elon lages lima
curso
de análise
vol. 2
PROJETO
EUCLIDES

curso de análise volume 2

elon lages lima

curso de análise volume 2

INSTITUTO NACIONAL DE MATEMÁTICA PURA E APLICADA

Impresso no Brasil / Printed in Brazil

Capa: Gian Calvi Criação Visual Ltda. / Sérgio R. Vaz

Títulos Publicados:

- Curso de Análise, Volume 1 - *Elon Lages Lima*
- Medida e Integração - *Pedro Jesus Fernandez*
- Aplicações da Topologia à Análise - *Chaim Samuel Hönig*
- Espaços Métricos - *Elon Lages Lima*
- Análise de Fourier e Equações Diferenciais Parciais - *Djairo Guedes de Figueiredo*
- Introdução aos Sistemas Dinâmicos - *Jacob Palis Junior e Wellington C. de Melo*
- Introdução à Álgebra - *Adilson Gonçalves*
- Aspectos Teóricos da Computação - *Cláudio L. Lucchesi, Imre Simon, Istvan Simon, Janos Simon e Tomasz Kowaltowski*
- Teoria Geométrica das Folheações - *Alcides Lins Neto e César Camacho*
- Geometria Riemanniana - *Manfredo P. do Carmo*
- Lições de Equações Diferenciais Ordinárias - *Jorge Sotomayor*
- Probabilidade: Um Curso em Nível Intermediário - *Barry R. James*
- Curso de Análise, Volume 2 - *Elon Lages Lima*
- Teoria Ergódica - *Ricardo Mañé*
- Teoria dos Números Algébricos - *Otto Endler*
- Operadores Auto-Adjuntos e Equações Diferenciais Parciais - *Javier Thayer*
- Equações Diferenciais Parciais: Uma Introdução - *Rafael Iório Jr. e Valéria Iório*
- Álgebra: Um Curso de Introdução - *Arnaldo Leite P. Garcia e Yves Albert E. Lequain*
- Grupo Fundamental e Espaços de Recobrimento - *Elon Lages Lima*
- Funções de uma Variável Complexa - *Alcides Lins Neto*
- Elementos de Álgebra - *Arnaldo Garcia e Yves Lequain*
- Introdução à Geometria Analítica Complexa - *Marcos Sebastiani*
- Curso de Teoria da Medida - *Augusto Armando de Castro Júnior*
- Introdução à Teoria da Medida - *Carlos Isnard*
- Introdução à Teoria de Controle e Programação Dinâmica - *Johann Baumeister e Antonio Leitão*
- Homologia Básica - *Elon Lages Lima*
- Teoria dos Números: um Passeio com Primos e outros Números Familiares pelo Mundo Inteiro - *Fabio Brochero Martinez, Carlos Gustavo Moreira, Nicolau Saldanha e Eduardo Tengan*
- Introdução à Análise Funcional – *César R. de Oliveira*

Distribuição:
IMPA
Estrada Dona Castorina, 110
22460-320 Rio de Janeiro, RJ
e-mail: ddic@impa.br
http://www.impa.br

Prefácio da primeira edição

Que outrem possa louvar esforço alheio,
Cousa é que se costuma e se deseja;
Mas louvar os meus próprios, arreceio
Que louvor tão suspeito mal me esteja;
E, para dizer tudo, temo e creio
Que qualquer longo tempo curto seja;
Mas, pois o mandas, tudo se te deve;
Irei contra o que devo e serei breve.

L. Camões "Os Lusíadas", Canto III.

Este segundo volume do Curso de Análise trata das funções de diversas variáveis reais. Como é natural, sua leitura pressupõe uma certa familiaridade com funções de uma variável. Além disso, admitem-se conhecidas algumas noções básicas de Álgebra Linear. O primeiro volume do Curso é a referência óbvia (mas não necessária) para os pré-requisitos de Análise. Para a Álgebra Linear, basta a leitura dos primeiros capítulos de qualquer dos bons livros existentes na praça.

As principais diferenças entre a Análise de uma e a de n variáveis têm suas origens em dois fatos. O primeiro é que a topologia dos subconjuntos de $\mathbb{R}^n$ fica muito mais complicada quando $n > 1$. (Por exemplo, os únicos subconjuntos conexos da reta são os intervalos, mas é impossível classificar topologicamente os subconjuntos conexos de $\mathbb{R}^n$ se $n \geq 2$.) Em segundo lugar, a Álgebra Linear, que em dimensão 1 é desnecessária, (pois as matrizes 1×1 e as transformações lineares de $\mathbb{R}$ em $\mathbb{R}$ se confundem com números reais) torna-se indispensável para formular os conceitos e demonstrar os teoremas do Cálculo Diferencial das funções com mais de uma variável.

A infinita variedade de tipos topológicos de subconjuntos do espaço $\mathbb{R}^n$, que servirão de domínios para as funções que estudaremos, empresta a este livro um conteúdo bastante geométrico, em contraste com o caráter predominantemente aritmético do Volume 1, e nos leva a dedicar todo o primeiro capítulo à topologia do espaço euclidiano. Os conceitos ali apresentados e o ponto de vista adotado situam-se no contexto da Topologia Geral. (No espírito de [8], por exemplo.) Salvo a

circunstância de serem mais complexas em dimensão maior do que um, estas noções não diferem muito das suas homônimas em dimensão 1.

Entretanto, a presença da Topologia se faz mais conspícua e profunda nos Capítulo IV e VII, onde surgem de modo natural as questões que levaram Gauss, Riemann, Kronecker e outros a abordarem problemas de Análise com métodos que mais tarde viriam a fazer parte da Topologia Diferencial. Seria inútil tentar evitar perguntas naturais de Análise, como as consideradas naqueles capítulos, sob a alegação de que as respostas envolvem conceitos não-triviais de Topologia, tais como homotopia, grau, etc. A atitude mais adequada nos parece a de enfrentar os problemas com as ferramentas criadas para esses fins, as quais posteriormente se difundiram e foram definitivamente incorporadas à Matemática de hoje.

Acreditamos que a exposição aqui feita das integrais curvilíneas e de superfície é suficientemente elementar para ser acessível a quem leu o primeiro volume e tenha alguma experiência com Álgebra Linear. De resto, é bom lembrar que esses assuntos não são novos em livros de Análise. A integral de Kronecker, por exemplo, já aparece no primeiro volume do "Traité d'Analyse" de Picard, em 1891.

Quanto à Álgebra Linear, sua posição no ensino de hoje é bem melhor do que era para os estudantes da minha geração. Não foi necessário incluir no livro uma introdução a essa disciplina, pois ela já se encontra suficientemente difundida, tanto na literatura disponível (mesmo em língua portuguesa) como em vários cursos em nível de graduação. É de se esperar que vetores, matrizes, transformações lineares, etc. constituam a linguagem natural para tratar o Cálculo Diferencial pois, afinal de contas, este se baseia na idéia de aproximar, na vizinhança de cada ponto de seu domínio, uma função "arbitrária" por uma função linear (chamada sua derivada) e, a partir das propriedades desta, (presumivelmente mais fáceis de constatar) obter informações sobre aquela.

Apenas a circunstância de que, em dimensão 1, uma transformação linear se confunde com um número real é que leva a considerar a derivada como sendo um número. Em dimensões superiores a verdade se revela: a derivada é uma transformação linear. Com base nesta observação, os resultados e métodos da Álgebra Linear são ubíquos nos tratamentos modernos do Cálculo Diferencial. Seu uso sistemático traz grandes vantagens de conceituação, notação, unificação e generalização.

Menos divulgada do que a Álgebra Linear propriamente dita, é seu

prolongamento conhecido como Álgebra Exterior. Esta fornece o formalismo adequado para o tratamento algébrico de certas noções geométricas elusivas, como a orientação, e o fundamento para o estudo das formas diferenciais. Como se sabe desde Elie Cartan, as formas constituem os objetos mais convenientes para serem colocados sob o sinal de integral, principalmente quando não se deseja assumir compromisso com algum sistema de coordenadas.

Admitindo conhecidas as propriedades elementares dos determinantes, começamos o último capítulo deste livro com uma exposição sucinta de Álgebra Exterior, a qual usamos como base para o estudo das integrais de superfície. Procuramos manter a generalidade e o formalismo dentro de limites aceitáveis enquanto que, por outro lado, exploramos algumas das excelsas virtudes do Cálculo Diferencial Exterior. Esperamos que o leitor possa sentir a elegância e a força dos métodos ali esboçados e se aposse desse poderoso instrumento para usos posteriores.

Ainda que isto não pareça claro nas páginas que se seguem, ao escrevê-las fui muito influenciado pelos livros de Picard, Courant, Phillips, Valiron e Graves citados na lista de referências. Cada um deles tem sido para mim um professor solícito, um conselheiro sábio, uma fonte de inspiração, um modelo a imitar.

Em termos mais concretos, fui também bastante estimulado por um grupo de amigos, temporariamente meus alunos num curso do IMPA, cujo entusiasmo pelo assunto e interesse pelo prosseguimento do livro foram uma importante ajuda. Não podendo mencionar a todos, gostaria pelo menos de destacar os que mais fiscalizaram meus descuidos e me auxiliaram na correção do manuscrito e das provas. Deixo consignados, pois, meus agradecimentos a Jonas de Miranda Gomes, José Felipe Voloch, Katia Frensel e Paulo Ney de Souza, aos quais deve também o leitor a ausência de alguns cochilos desagradáveis.

No plano da elaboração gráfica, sou grato a Wilson Góes, pela boa vontade e eficiência no preparo de mais um dos inúmeros manuscritos que já lhe entreguei. Agradeço ainda aos senhores João D. Affonso e Wilson Siviero, da AM Produções Gráficas Ltda. e Gráfica Editora Hamburg Ltda., respectivamente, e às excelentes equipes que eles comandam, pela competência, pela boa vontade e pelo excepcional acolhimento que têm dado às nossas publicações.

Rio de Janeiro, 14 de julho de 1981
Elon Lages Lima

Prefácio da décima edição

As sucessivas edições deste livro se beneficiaram da colaboração de leitores atentos, que apontaram várias correções a serem feitas. Dentre eles, quero agradecer a Florêncio Guimarães, Pedro Zühke de Oliveira, Diógenes Justo e Rick Rischter.

A oitava edição foi redigitada eletronicamente por Wilson Góes, e as figuras executadas por Francisco Petrúcio Cavalcante. Rogério Dias Trindade encarregou-se de incorporar as revisões. A todos esses amigos quero agradecer cordialmente.

Por último, mas não por menos, registro aqui minha gratridão a todos os colegas que adotaram o livro em seus cursos e aos leitores que nele estudaram.

Rio de Janeiro, dezembro de 2007

Elon Lages Lima

Conteúdo

Capítulo I

Topologia do Espaço Euclidiano

1 O espaço vetorial $\mathbb{R}^n$

Seja n um número natural. O espaço *euclidiano n-dimensional* é o produto cartesiano de n fatores iguais a $\mathbb{R}$:

$$\mathbb{R}^n = \mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}.$$

Os pontos de $\mathbb{R}^n$ são, pois, todas as n-listas $x = (x_1, \ldots, x_n)$ cujas coordenadas $x_1, \ldots, x_n$ são números reais. Dados $x = (x_1, \ldots, x_n)$ e $y = (y_1, \ldots, y_n)$ em $\mathbb{R}^n$ tem-se $x = y$ se, e somente se, $x_1 = y_1, \ldots, x_n = y_n$.

$\mathbb{R}^1 = \mathbb{R}$ é a *reta*, isto é, o conjunto dos números reais. $\mathbb{R}^2$ é o *plano*, ou seja, o conjunto dos pares ordenados $z = (x, y)$ de números reais. $\mathbb{R}^3$ é o espaço euclidiano tridimensional da geometria analítica, cujos pontos são os ternos ordenados $p = (x, y, z)$. Às vezes é conveniente considerar também $\mathbb{R}^0 = \{0\}$, o "espaço de dimensão zero", formado pelo único ponto 0.

Dados $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$ em $\mathbb{R}^n$ e um número real α, definimos a *soma* $x + y$ e o *produto* $\alpha \cdot x$ pondo

$$x + y = (x_1 + y_1, \ldots, x_n + y_n),$$
$$\alpha \cdot x = (\alpha x_1, \ldots, \alpha x_n).$$

Estas operações fazem de $\mathbb{R}^n$ um espaço vetorial de dimensão n sobre o corpo dos reais, no qual o elemento neutro para a adição é $0 = (0, \ldots, 0)$ e o simétrico de $x = (x_1, \ldots, x_n)$ é $-x = (-x_1, \ldots, -x_n)$.

Os elementos de $\mathbb{R}^n$ serão às vezes chamados *pontos* e às vezes *vetores*. Geometricamente, considerar $x \in \mathbb{R}^n$ como vetor significa imaginar a seta que tem origem no ponto 0 e extremidade em x.

No espaço vetorial $\mathbb{R}^n$, destaca-se a *base canônica* ou *base natural* $\{e_1, \ldots, e_n\}$, formada pelos vetores

$$e_1 = (1, 0, \ldots, 0),\ e_2 = (0, 1, \ldots, 0), \ldots, e_n = (0, \ldots, 0, 1),$$

que têm uma coordenada igual a 1 e as outras nulas. Dado $x = (x_1, \ldots, x_n)$ em $\mathbb{R}^n$, tem-se

$$x = x_1 \cdot e_1 + \cdots + x_n \cdot e_n\,.$$

A base canônica do espaço euclidiano permite estabelecer uma bijeção natural entre o conjunto $\mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$ das aplicações (ou transformações) lineares $A\colon \mathbb{R}^m \to \mathbb{R}^n$ e o conjunto $M(n \times m)$ das matrizes reais (a_{ij}) com n linhas e m colunas. A matriz (a_{ij}) que corresponde à transformação linear A é definida pelas igualdades

$$(*) \qquad A \cdot e_j = \sum_{i=1}^{n} a_{ij} e_i \quad (j = 1, \ldots, m).$$

Assim, a matriz (a_{ij}) da transformação linear $A\colon \mathbb{R}^m \to \mathbb{R}^n$ tem como colunas os m vetores $A \cdot e_j = (a_{1j}, \ldots, a_{nj}) \in \mathbb{R}^n$, imagens (ou transformados) por A dos vetores da base canônica de $\mathbb{R}^m$. Reciprocamente, dada uma matriz (a_{ij}) com n linhas e m colunas, as igualdades (*) definem os valores de uma aplicação linear $A\colon \mathbb{R}^m \to \mathbb{R}^n$ nos m vetores da base canônica. Como se sabe, isto é suficiente para definir o valor de A em qualquer vetor $x = (x_1, \ldots, x_m) \in \mathbb{R}^m$, tendo-se $A \cdot x = x_1 \cdot Ae_1 + \cdots + x_m \cdot Ae_m\,.$

Cada matriz real $n \times m$ pode ser considerada como um ponto do espaço euclidiano $\mathbb{R}^{nm}$: basta escrever suas colunas, uma após a outra, numa linha. Assim, sempre que for conveniente, podemos substituir o conjunto $\mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$ das aplicações lineares de $\mathbb{R}^m$ em $\mathbb{R}^n$, ora pelo conjunto $M(n \times m)$ das matrizes reais com n linhas e m colunas, ora pelo espaço euclidiano nm-dimensional $\mathbb{R}^{nm}$.

Um tipo especialmente simples de aplicações lineares é fornecido pelos *funcionais lineares* $f\colon \mathbb{R}^m \to \mathbb{R}$. Dado o funcional linear f, sejam $a_1 = f(e_1), \ldots, a_m = f(e_m)$ os valores que ele assume nos vetores da base

canônica. Para qualquer $x = (x_1, \ldots, x_m) \in \mathbb{R}^m$, temos $x = \Sigma\, x_i e_i$, logo $f(x) = \Sigma\, x_i f(e_i)$, ou seja:

$$f(x) = a_1 x_1 + \cdots + a_m x_m\,.$$

Note que $(a_1, \ldots, a_m)$ é a matriz $1 \times m$ da aplicação linear f.

Em particular, dado o número natural $i \in [1, m]$, seja $\pi_i \colon \mathbb{R}^m \to \mathbb{R}$ o funcional que se anula em todos os vetores da base canônica exceto um, a saber, o vetor e_i, para o qual se tem $\pi_i(e_i) = 1$. Então $\pi_i(x) = x_i = i$-ésima coordenada do ponto $x \in \mathbb{R}^m$. Assim, π_i é a *i-ésima projeção* do produto cartesiano $\mathbb{R}^m$ no seu fator $\mathbb{R}$. Os funcionais lineares $\pi_1, \ldots, \pi_m$ constituem uma base do espaço vetorial $\mathcal{L}(\mathbb{R}^m; \mathbb{R}) = (\mathbb{R}^m)^*$, chamada a *base dual* da base canônica de $\mathbb{R}^m$.

Uma aplicação $\varphi \colon \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^p$ chama-se *bilinear* quando é linear separadamente em relação a cada uma das suas variáveis. Isto significa que se tem:

$$\begin{aligned} \varphi(x + x', y) &= \varphi(x, y) + \varphi(x', y); \\ \varphi(x, y + y') &= \varphi(x, y) + \varphi(x, y'); \\ \varphi(\alpha x, y) &= \alpha \cdot \varphi(x, y); \\ \varphi(x, \alpha y) &= \alpha \cdot \varphi(x, y), \end{aligned}$$

quaisquer que sejam $x, x' \in \mathbb{R}^m$, $y, y' \in \mathbb{R}^n$ e $\alpha \in \mathbb{R}$. Se φ é bilinear então, para $x = (x_1, \ldots, x_m)$ e $y = (y_1, \ldots, y_n)$ arbitrários vale: $\varphi(x, y) = \varphi(\Sigma\, x_i e_i, \Sigma\, y_j e_j) = \sum_{i,j} x_i y_j\, \varphi(e_i, e_j)$, de modo que φ fica inteiramente determinada pelos mn valores $\varphi(e_i, e_j) \in \mathbb{R}^p$ que assume nos pares ordenados de vetores básicos (e_i, e_j). Note que $\varphi(x, 0) = \varphi(0, y) = 0$ quaisquer que sejam $x \in \mathbb{R}^m$, $y \in \mathbb{R}^n$.

2 Produto interno e norma

Um *produto interno* num espaço vetorial real E é uma regra que faz corresponder a cada par de vetores $x, y \in E$ um número real, indicado por $\langle x, y \rangle$, de tal modo que, para quaisquer x, x', $y \in E$ e $\alpha \in \mathbb{R}$, se

tenham:

PI.1 $\langle x, y\rangle = \langle y, x\rangle$;
PI.2 $\langle x + x', y\rangle = \langle x, y\rangle + \langle x', y\rangle$;
PI.3 $\langle \alpha x, y\rangle = \alpha \cdot \langle x, y\rangle = \langle x, \alpha y\rangle$;
PI.4 $x \neq 0 \Rightarrow \langle x, x\rangle > 0$.

Isto se exprime dizendo que um produto interno é uma função real simétrica, bilinear, positiva definida, $E \times E \to \mathbb{R}$.

O exemplo mais importante é o *produto interno canônico* do espaço euclidiano $\mathbb{R}^n$, o qual é dado por

$$\langle x, y\rangle = x_1y_1 + \cdots + x_ny_n\,,$$

onde $x = (x_1, \ldots, x_n)$ e $y = (y_1, \ldots, y_n)$.

Salvo menção explícita em contrário, o produto interno canônico é o único que consideraremos no espaço euclidiano $\mathbb{R}^n$.

Observação: A maneira mais geral de se definir um produto interno em $\mathbb{R}^n$ é a seguinte: toma-se uma matriz real (a_{ij}), $n \times n$, simétrica, positiva (isto é, $\Sigma\, a_{ij}x_ix_j > 0$ para todo $x \neq 0$), e põe-se $\langle x, y\rangle = \sum_{i,j=1}^{n} a_{ij}x_iy_j$. O produto interno canônico corresponde a tomar a matriz identidade.

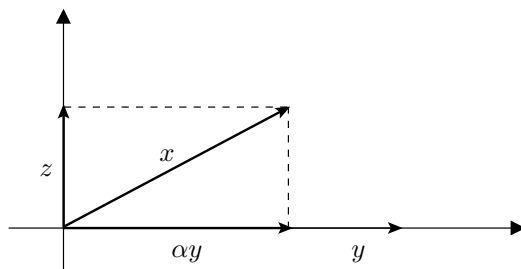
Dado $x \in \mathbb{R}^n$, escreveremos $|x| = \sqrt{\langle x, x\rangle}$, ou seja,

$$|x| = \sqrt{x_1^2 + \cdots + x_n^2}.$$

Tem-se $|x|^2 = \langle x, x\rangle$, de modo que $|x| = 0 \Leftrightarrow x = 0$ e $|x| > 0 \Leftrightarrow x \neq 0$.

O número $|x|$ chama-se a *norma euclidiana* ou o *comprimento* do vetor $x \in \mathbb{R}^n$.

Dois vetores $x, y \in \mathbb{R}^n$ dizem-se *ortogonais* quando $\langle x,y\rangle = 0$. Evidentemente, 0 é ortogonal a todos os vetores de $\mathbb{R}^n$. Também e_i é ortogonal a e_j se $i \neq j$.

(2.1) Um caso menos banal de ortogonalidade é o seguinte: dados $x, y \in \mathbb{R}^n$, com $y \neq 0$, e pondo-se $\alpha = \langle x, y\rangle/|y|^2$, o vetor $z = x - \alpha y$ é ortogonal a y.

Com efeito, $\langle z, y\rangle = \langle x - \alpha y, y\rangle = \langle x, y\rangle - \alpha\langle y, y\rangle = \langle x, y\rangle - \alpha|y|^2 = \langle x, y\rangle - \langle x, y\rangle = 0$.

Usaremos agora a observação acima para demonstrar a desigualdade fundamental da Geometria Euclidiana.

Teorema 1 (Desigualdade de Cauchy-Schwarz). *Para quaisquer $x, y \in \mathbb{R}^n$, tem-se $|\langle x, y\rangle| \leq |x| \cdot |y|$. Vale a igualdade se, e somente se, um dos vetores x, y é um múltiplo escalar do outro.*

Demonstração: Isto é óbvio se $y = 0$. Se, porém, for $y \neq 0$, poremos $\alpha = \langle x, y\rangle/|y|^2$. Como acabamos de ver, o vetor $z = x - \alpha y$ é ortogonal a y. Segue-se daí que

$$|x|^2 = \langle z + \alpha y, z + \alpha y\rangle = |z|^2 + \alpha^2|y|^2, \quad \text{donde}$$
$$|x|^2 \geq \alpha^2|y|^2 = \langle x, y\rangle^2/|y|^2, \quad \text{ou seja:}$$

$|x|^2|y|^2 \geq \langle x, y\rangle^2$, como queríamos demonstrar. Vale a igualdade se, e somente se, $z = 0$, ou seja, $x = \alpha \cdot y$.

Observação: A prova acima é válida para qualquer produto interno, com $|x| = \sqrt{\langle x, x\rangle}$.

A norma euclidiana $|x| = \sqrt{\langle x, x\rangle}$ goza das seguintes propriedades, onde $x, y \in \mathbb{R}^n$, $\alpha \in \mathbb{R}$ e $|\alpha|$ significa o valor absoluto do número real α:

N1. $|x + y| \leq |x| + |y|$;
N2. $|\alpha \cdot x| = |\alpha| \cdot |x|$;
N3. $x \neq 0 \Rightarrow |x| > 0$.

As duas últimas são evidentes e a primeira propriedade decorre da desigualdade de Cauchy-Schwarz. Com efeito:

$$|x + y|^2 = \langle x + y, x + y\rangle = |x|^2 + |y|^2 + 2\langle x, y\rangle \leq$$
$$\leq |x|^2 + |y|^2 + 2|x| \cdot |y| = \big(|x| + |y|\big)^2.$$

De um modo geral, uma *norma* num espaço vetorial E é qualquer função real $|\ |: E \to \mathbb{R}$ que cumpra as condições N1, N2 e N3 acima estipuladas.

Note-se que, tomando $\alpha = 0$ em N2, vem $|0| = 0$ e pondo $\alpha = -1$ vem $|-x| = |x|$. Pondo $y = -x$ em N1 obtemos $0 = |0| = |x + (-x)| \leq |x| + |-x| = 2|x|$, donde $|x| \geq 0$ para todo $x \in E$.

Observe que $|x - y| = |-(x - y)| = |y - x|$.

Há uma infinidade de normas que se podem considerar no espaço euclidiano $\mathbb{R}^n$. A norma euclidiana é motivada pela fórmula do comprimento de um vetor no plano em coordenadas cartesianas, que se prova com o Teorema de Pitágoras. Para noções geométricas, ela é a mais natural. Quando não dissermos explicitamente qual a norma que estamos considerando em $\mathbb{R}^n$, fica subentendido que se trata da euclidiana. Por outro lado, há duas normas que são de manipulação formal mais simples, as quais poderemos utilizar em $\mathbb{R}^n$, quando houver conveniência. Elas são:

$$|x|_M = \text{ Máx. } \{|x_1|, \ldots, |x_n|\}, \qquad \text{(Norma do Máximo)}$$
$$|x|_S = |x_1| + \cdots + |x_n|. \qquad \text{(Norma da Soma)}$$

As condições N1, N2 e N3 são fáceis de verificar para estas duas normas. Também é simples mostrar que, para todo $x \in \mathbb{R}^n$ vale

$$|x|_M \leq |x| \leq |x|_S \leq n \cdot |x|_M\,,$$

onde $|x|$ é a norma euclidiana. Estas desigualdades servirão para mostrar que as 3 normas acima são equivalentes, no sentido que definiremos no §4 a seguir.

Advertimos que as notações $|x|_M$ e $|x|_S$ não serão usadas sistematicamente para indicar as normas que acabamos de definir. Assim, por exemplo, se estivermos aplicando apenas a norma $|x|_M$ num determinado contexto, poderemos indicá-la com a notação $|x|$, por simplicidade.

Uma norma num espaço vetorial E dá origem a uma noção de distância em E. Dados $x, y \in E$, a *distância* de x a y é definida por

$$d(x, y) = |x - y|.$$

As condições N1, N2 e N3, que a norma satisfaz, implicam imediatamente que a distância goza das seguintes propriedades, para $x, y, z \in E$

quaisquer:

$$\begin{aligned}
&\text{d1.} \quad d(x,z) \le d(x,y) + d(y,z);\\
&\text{d2.} \quad d(x,y) = d(y,x);\\
&\text{d3.} \quad x \ne y \Rightarrow d(x,y) > 0.
\end{aligned}$$

A primeira dessas propriedades é chamada a *desigualdade do triângulo.* Ela exprime que o comprimento de um dos lados de um triângulo não excede a soma dos comprimentos dos outros dois lados. Sua demonstração se faz assim:

$$\begin{aligned}
d(x,z) = |x-z| = |(x-y)+(y-z)| \le |x-y| + |y-z| =\\
= d(x,y) + d(y,z).
\end{aligned}$$

Uma norma arbitrária $|\ |$ num espaço vetorial E pode não provir de um produto interno, isto é, nem sempre existe um produto interno $\langle\ ,\ \rangle$ em E tal que $|x|^2 = \langle x, x\rangle$ para todo $x \in E$. Com efeito, se a norma provém de um produto interno, então vale a *identidade do paralelogramo*

$$|x+y|^2 + |x-y|^2 = 2\big(|x|^2 + |y|^2\big),$$

que diz ser a soma dos quadrados das diagonais de um paralelogramo igual à soma dos quadrados dos seus quatro lados. Com efeito, temos

$$\begin{aligned}
|x+y|^2 = \langle x+y, x+y\rangle = |x|^2 + |y|^2 + 2\langle x,y\rangle,\\
|x-y|^2 = \langle x-y, x-y\rangle = |x|^2 + |y|^2 - 2\langle x,y\rangle.
\end{aligned}$$

Somando membro a membro, obtemos a identidade desejada.

Ora, a identidade do paralelogramo não é válida para toda norma. Por exemplo, as normas $|x|_M = \text{Máx}\,\{|x_1|, \ldots, |x_n|\}$ e $|x|_S = \Sigma|x_i|$ em $\mathbb{R}^n$ não a cumprem. (Tome $x = e_1$ e $y = e_2$.) Logo essas normas não provêm de produto interno algum em $\mathbb{R}^n$.

Uma conseqüência simples, porém útil, dos axiomas N1, N2 e N3 é que, dada uma norma $|\ |$ em $\mathbb{R}^n$, para $x, y \in \mathbb{R}^n$ quaisquer tem-se:

$$\big|\,|x| - |y|\,\big| \le |x-y|.$$

Com efeito, a desigualdade acima significa que $-|x-y| \le |x| - |y| \le |x-y|$, ou seja, que $|y| \le |x| + |x-y|$ e $|x| \le |y| + |x-y|$. Isto se segue de N1, pois $y = x + (y-x)$ e $x = y + (x-y)$.

3 Números complexos

Além da estrutura de espaço vetorial, existente em todos os espaços euclidianos $\mathbb{R}^n$, o plano $\mathbb{R}^2$ possui uma multiplicação, segundo a qual o produto de $z = (x, y)$ por $w = (u, v)$ é definido pela seguinte regra:

$$z \cdot w = (xu - yv, xv + yu).$$

Esta multiplicação é comutativa, associativa, distributiva em relação à adição e admite o vetor $e_1 = (1, 0)$ como elemento neutro multiplicativo, isto é, tem-se $z \cdot e_1 = z$ para todo $z \in \mathbb{R}^2$.

Quando quisermos dizer que estamos considerando $\mathbb{R}^2$ munido da multiplicação acima definida, usaremos a notação $\mathbb{C} = \mathbb{R}^2$ e chamaremos os pontos $z = (x, y)$ de *números complexos*.

Como $(x, 0) = (y, 0) = (x + y, 0)$ e, na multiplicação de complexos, vale $(x, 0) \cdot (y, 0) = (xy, 0)$, concluimos que a imersão natural $f : \mathbb{R} \to \mathbb{C}$, definida por $f(x) = (x, 0)$, preserva as operações de adição e multiplicação, isto é, $f(x + y) = f(x) + f(y)$ e $f(x \cdot y) = f(x) \cdot f(y)$. Em vista disto, identificaremos x com $f(x)$, isto é, consideraremos $\mathbb{R} \subset \mathbb{C}$, $(x, 0) = x$ e, em particular, $(0, 0) = 0$, $e_1 = 1$ (unidade real). Usa-se a notação $i = (0, 1) = e_2$ (unidade imaginária). Como $i^2 = (-1, 0) = -1$, escreve-se às vezes $i = \sqrt{-1}$. Para todo número complexo $z = (x, y)$, temos $z = xe_1 + ye_2 = x + yi$. Com esta notação, a definição do produto $z \cdot w$ se torna natural: para multiplicar $z = x + yi$ por $w = u + vi$, opera-se como se fosse com números reais, tendo-se apenas o cuidado de substituir i^2 por -1. Assim:

$$\begin{aligned} z \cdot w = (x + yi)(u + vi) = xu + (xv + yu)i + yv \cdot i^2 = \\ = (xu - yv) + (xv + yu)i. \end{aligned}$$

No número complexo $z = x + yi$, x chama-se a *parte real* e y a *parte imaginária*. O número $\bar{z} = x - yi$ chama-se o *conjugado* de z. Por analogia com os reais, a norma euclidiana (única a ser usada) $|z| = \sqrt{x^2 + y^2}$ chama-se o *módulo* do número complexo z. Verifica-se facilmente que $|z| = |\bar{z}|$, $z \cdot \bar{z} = x^2 + y^2 = |z|^2$ e $|z \cdot w| = |z| \cdot |w|$ para quaisquer $z, w \in \mathbb{C}$.

Se $z \neq 0$, pondo $w = \bar{z}/|z|^2$ obtemos $z \cdot w = 1$. Logo todo elemento $z = (x, y) \neq 0$ em $\mathbb{C}$ possui o inverso multiplicativo

$$z^{-1} = \frac{\bar{z}}{|z|^2} = \left(\frac{x}{x^2 + y^2}, \frac{-y}{x^2 + y^2} \right).$$

Em particular, se $|z| = 1$ então $z^{-1} = \bar{z}$.

O conjunto $\mathbb{C}$ dos números complexos é portanto um *corpo*. (Vide Cap. 3 do Volume 1.) O conjunto $\mathbb{C} - \{0\}$, dos números complexos não-nulos, é um grupo abeliano multiplicativo. O círculo unitário $S^1 = \{x + yi \in \mathbb{C}; x^2 + y^2 = 1\}$, conjunto dos números complexos de módulo 1, é um subgrupo multiplicativo de $\mathbb{C} - \{0\}$, pois se $z, w \in S^1$ então $|z \cdot w| = |z| \cdot |w| = 1$ e $|z^{-1}| = |\bar{z}| = 1$, logo $z \cdot w,\ \ z^{-1} \in S^1$.

Definiremos a *aplicação exponencial* $\xi\colon \mathbb{R} \to S^1$, pondo $\xi(t) = \cos t + i \cdot \operatorname{sen} t$.

(3.1) *ξ é um homomorfismo do grupo aditivo $\mathbb{R}$ dos números reais sobre o grupo multiplicativo S^1 dos números complexos de módulo* 1.

Demonstração: Para quaisquer $s, t \in \mathbb{R}$, temos $\xi(s+t) = \cos(s+t) + i \cdot \operatorname{sen}(s+t) = (\cos s \cdot \cos t - \operatorname{sen} s \cdot \operatorname{sen} t) + i(\cos s \cdot \operatorname{sen} t + \operatorname{sen} s \cdot \cos t) = (\cos s + i \operatorname{sen} s)(\cos t + i \operatorname{sen} t) = \xi(s) \cdot \xi(t)$. (Vide Volume 1, pág. 392.) Logo ξ é um homomorfismo. Para provar que ξ é sobrejetiva, seja $x + yi \in S^1$, isto é, $x^2 + y^2 = 1$. Então $x \in [-1, 1]$. Como $\cos\colon \mathbb{R} \to [-1, 1]$ é sobrejetiva, existe $t \in \mathbb{R}$ tal que $\cos t = x$. Logo $\operatorname{sen} t = \pm\sqrt{1 - \cos^2 t} = \pm\sqrt{1 - x^2} = \pm y$. Se for $\operatorname{sen} t = y$, então $\xi(t) = x + yi$. Se for $\operatorname{sen} t = -y$ então $\xi(-t) = \cos(-t) + i \operatorname{sen}(-t) = \cos t - i \operatorname{sen} t = x + yi$. De qualquer modo, ξ é sobrejetiva.

Tem-se $\xi(s) = \xi(t)$ se, e somente se, os números reais s, t têm o mesmo cosseno e o mesmo seno, donde $s - t = 2k\pi,\ \ k \in \mathbb{Z}$. Em particular, $\xi(t) = 1$, se, e somente se $t = 2k\pi,\ \ k \in \mathbb{Z}$, ou seja, o núcleo do homomorfismo ξ é o conjunto dos múltiplos inteiros de 2π.

Desenvolvendo $\cos t$ e $\operatorname{sen} t$ em série de potências, obtemos a igualdade formal $\cos t + i \operatorname{sen} t = \sum_{n=0}^{\infty} (it)^n/n!$. Isto sugeriu a Euler a notação $e^{it} = \cos t + i \operatorname{sen} t$ para o que chamamos acima $\xi(t)$. O fato de $\xi\colon \mathbb{R} \to S^1$ ser um homomorfismo significa que $e^{i(s+t)} = e^{is} \cdot e^{it},\ \ (e^{it})^{-1} = e^{-it} =$ conjugado de e^{it}. Mais geralmente, para todo $n \in \mathbb{Z}$, vale $(e^{it})^n = e^{int}$. Isto significa que $(\cos t + i \operatorname{sen} t)^n = \cos(nt) + i \operatorname{sen}(nt)$, igualdade conhecida como *fórmula de Moivre*.

Todo número complexo $z \neq 0$ é o produto $z = |z| \cdot z/|z|$ de um número real positivo $|z| = r$ por um número complexo $z/|z| = e^{it}$, de módulo 1. A expressão $z = r \cdot e^{it}$, com $r > 0$ e $t \in \mathbb{R}$, chama-se a *forma polar* do número complexo $z \neq 0$. Nela, o módulo $r = |z|$ é univocamente determinado a partir de z, mas o número real t, chamado (um) *argumento* de z, não é único: se t é argumento de z então os demais

argumentos são os números $t + 2k\pi, \quad k \in \mathbb{Z}$.

A forma polar é conveniente para a multiplicação de números complexos: se $z = r \cdot e^{it}$ e $w = q \cdot e^{is}$ então $w \cdot z = qr \cdot e^{i(s+t)}$. Assim, para multiplicar dois números complexos, multiplicam-se os módulos e somam-se os argumentos. Para a divisão, temos $w \cdot z^{-1} = q \cdot r^{-1}\, e^{i(s-t)}$, ou seja, dividem-se os módulos e subtraem-se os argumentos.

Seja $A = a + bi \neq 0$. Em virtude da distributividade e da comutatividade da multiplicação de números complexos, a multiplicação por A define uma transformação linear $\widetilde{A}\colon \mathbb{R}^2 \to \mathbb{R}^2$, com $\widetilde{A} \cdot z = A \cdot z$. Note que $\widetilde{A}$ é invertível, sendo sua inversa a multiplicação por A^{-1}. A matriz da transformação $\widetilde{A}$ tem como vetores-coluna $\widetilde{A} \cdot e_1 = A \cdot 1 = (a, b)$ e $\widetilde{A} \cdot e_2 = A \cdot i = (-b, a)$, logo tal matriz tem a forma

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

O determinante desta matriz é igual a $a^2 + b^2 = |A|^2$.

Quando $|A| = 1$ então $|A \cdot z| = |z|$ e $\langle Aw, Az\rangle = \langle w, z\rangle$ para quaisquer $w, z \in \mathbb{R}^2$, como se verifica diretamente. Neste caso, a transformação linear $\widetilde{A}$ preserva a norma e o produto interno em $\mathbb{R}^2$. Pondo então $A = \cos\theta + i \operatorname{sen}\theta$, a matriz de $\widetilde{A}$ tem a forma

$$\begin{pmatrix} cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}$$

e $\widetilde{A}$ se chama uma *rotação positiva* (pois $\det \widetilde{A} > 0$).

Sejam $u, v, w, z \in S^1$ vetores unitários no plano. Diremos que o *ângulo* $\measuredangle(u, v)$ *é igual ao ângulo* $\measuredangle(w, z)$ quando existir uma rotação positiva, isto é, um número complexo A de módulo 1 tal que $A \cdot u = w$, $A \cdot v = z$. Isto equivale a afirmar que $vu^{-1} = zw^{-1}$. Estamos definindo igualdade entre ângulos "orientados", isto é, $\measuredangle(v, u)$ se $v \neq \pm u$. Mais precisamente $\measuredangle(u, v) = \measuredangle(\bar{v}, \bar{u})$, pois $\bar{u} \cdot \bar{v}^{-1} = v \cdot u^{-1}$.

Mais geralmente, se $u, v, w, z \in \mathbb{C} - \{0\}$, *dizemos que* $\measuredangle(u, v) = \measuredangle(w, z)$ quando $\measuredangle\left(\dfrac{u}{|u|}, \dfrac{v}{|v|}\right) = \measuredangle\left(\dfrac{w}{|w|}, \dfrac{z}{|z|}\right)$, ou seja, quando os ângulos entre os vetores unitários correspondentes são iguais. Isto significa que existe uma rotação positiva $A \in S^1$ tal que $A \cdot u = \alpha \cdot w$, $A \cdot v = \beta \cdot z$ com $\alpha > 0$ e $\beta > 0$. (Aqui $\alpha = |u|/|w|$ e $\beta = |v|/|z|$.)

Tem-se $\measuredangle(u, v) = \measuredangle(w, z)$ se, e somente se, $v \cdot u^{-1} = q \cdot e^{it}$ e $z \cdot w^{-1} = r \cdot e^{it}$, com $q > 0, \quad r > 0$ e $t \in \mathbb{R}$. O número t se chama uma *medida*

em radianos do ângulo $\sphericalangle(w, z)$. As outras medidas desse ângulo são $t + 2k\pi, \quad k \in \mathbb{Z}$. Segue-se daí que são bem definidos o cosseno e o seno do ângulo $\sphericalangle(w, z)$: se $zw^{-1} = r \cdot e^{it}$ então pomos $\cos \sphericalangle(w, z) = \cos t$ e $\operatorname{sen} \sphericalangle(w, z) = \operatorname{sen} t$. Em outras palavras, se $w = q \cdot e^{is}$ e $z = r \cdot e^{it}$ então $cos \sphericalangle(w, z) = \cos(t - s)$ e $\operatorname{sen} \sphericalangle(w, z) = \operatorname{sen}(t - s)$. Observe que $\cos \sphericalangle(w, z) = \cos \sphericalangle(z, w)$ mas $\operatorname{sen} \sphericalangle(w, z) = -\operatorname{sen} \sphericalangle(z, w)$.

A transformação linear $\widetilde{A}\colon \mathbb{R}^2 \to \mathbb{R}^2$, definida por $\widetilde{A} \cdot z = A \cdot z$ (onde A é um número complexo $\neq 0$) preserva ângulos, ou seja, se w e z são vetores não nulos, $\sphericalangle(A \cdot w, A \cdot z) = \sphericalangle(w, z)$. $\widetilde{A}$ é o que se chama uma *semelhança positiva*: consta de uma rotação positiva (multiplicação por $A/|A|$) seguida de uma homotetia (multiplicação pelo número real $|A| > 0$).

Se $w = |w|(\cos s + i \operatorname{sen} s)$ e $z = |z|(\cos t + i \operatorname{sen} t)$ são vetores não-nulos no plano, seu produto interno é

$$\begin{aligned}\langle w, z \rangle = |w|\,|z|\,(\cos s \cdot \cos t + \operatorname{sen} s \cdot \operatorname{sen} t) = |w|\,|z|\,\cos(t - s) = \\ = |w|\,|z|\,\cos \sphericalangle(w, z),\end{aligned}$$

o que nos dá de volta a definição clássica do produto interno: o produto dos módulos vezes o cosseno do ângulo formado pelos dois vetores.

4 Bolas e conjuntos limitados

Uma norma em $\mathbb{R}^n$ permite que sejam definidas algumas noções geométricas básicas, que passamos a apresentar.

A *bola aberta* de centro num ponto $a \in \mathbb{R}^n$ e raio $r > 0$ é o conjunto dos pontos $x \in \mathbb{R}^n$ cuja distância ao ponto a é menor do que r. Usaremos a notação $B(a; r)$ para indicar esse conjunto. Assim:

$$B(a; r) = \{x \in \mathbb{R}^n; |x - a| < r\}.$$

Analogamente definiremos a *bola fechada* $B[a; r]$ e a *esfera* $S[a; r]$, ambas com centro a e raio r, pondo:

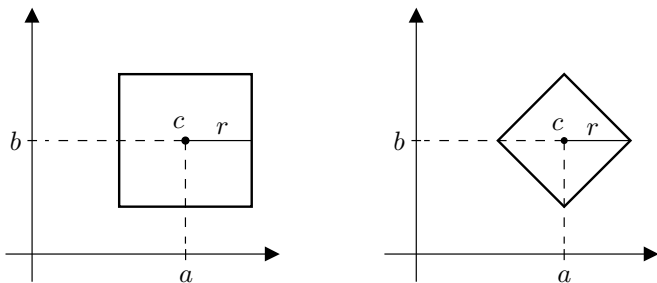
$$B[a; r] = \{x \in \mathbb{R}^n; |x - a| \leq r\} \text{ e } S[a; r] = \{x \in \mathbb{R}^n; |x - a| = r\}.$$

Segue-se que $B[a; r] = B(a; r) \cup S[a; r]$.

Quando $n = 1$, a bola aberta $B(a; r)$ de centro a e raio r na reta $\mathbb{R}$ é o intervalo aberto $(a - r, a + r)$, a bola fechada $B[a; r]$ é o intervalo fechado

$[a-r, a+r]$ e a esfera $S[a;r]$ reduz-se ao conjunto formado pelos dois pontos $a-r$ e $a+r$. Note que as três normas usuais do espaço euclidiano coincidem quando $n=1$.

Para $n=2$, tomando a norma euclidiana, as bolas no plano chamam-se *discos* (abertos ou fechados) e as esferas reduzem-se a círculos.

Se $n=3$, a norma euclidiana define no espaço bolas e esferas que correspondem às imagens que delas fazemos.

A forma geométrica das bolas e esferas dependem em geral, é claro, da norma que se usa. Se, em vez da norma euclidiana, tomarmos no plano $\mathbb{R}^2$, por exemplo, $|z|_M = \max\{|x|, |y|\}$ como norma de $z=(x,y)$, a bola de centro $c=(a,b)$ e raio r será o quadrado de lados paralelos aos eixos de coordenadas, cada lado tendo comprimento $2r$, e diagonais cortando-se no ponto c. Por outro lado, se tomarmos a norma da soma, $|z|_S = |x| + |y|$, a bola de centro c e raio r será o quadrado cujas diagonais são paralelas aos eixos coordenados, ambas com comprimento $2r$, encontrando-se ainda no ponto c.

Analogamente, a norma do máximo no espaço $\mathbb{R}^3$ produz bolas com forma de cubos de arestas paralelas aos eixos, enquanto que, na norma da soma, as bolas em $\mathbb{R}^3$ são octaedros com diagonais paralelas aos eixos.

De um modo mais geral, observamos que se $a=(a_1,\ldots,a_n)$ então a bola $B[a;r] \subset \mathbb{R}^n$, definida pela norma $|x|_M = \max\{|x_1|,\ldots,|x_n|\}$ é o produto cartesiano $B[a;r]{=}[a_1-r, a_1+r] \times \cdots \times [a_n-r, a_n+r]$. Com efeito, $|x-a| \le r \Leftrightarrow |x_1-a_1| \le r, \ldots, |x_n-a| \le r$. Do mesmo modo, usando ainda a norma do máximo, temos $B(a;r) = \prod_{i=1}^{n} (a_i - r, a_i + r)$ para a bola aberta. Estas propriedades tornam a norma do máximo conveniente em relação ao produto cartesiano.

O teorema seguinte diz que as bolas relativas a diferentes normas em $\mathbb{R}^n$ têm em comum o fato de serem convexas.

Sejam $x, y \in \mathbb{R}^n$. O *segmento de reta* de extremos x, y é o conjunto

$$[x, y] = \{(1-t)x + ty; 0 \leq t \leq 1\}.$$

Um subconjunto $X \subset \mathbb{R}^n$ diz-se *convexo* quando contém qualquer segmento de reta cujos extremos pertençam a X, ou seja: $x, y \in X \Rightarrow [x, y] \subset X$.

Todo subespaço vetorial $E \subset \mathbb{R}^n$ é convexo; toda *variedade afim* $a + E = \{a + x; x \in E\}$ (onde $E \subset \mathbb{R}^n$ é um subespaço vetorial) é convexa. Se $X \subset \mathbb{R}^m$ e $Y \subset \mathbb{R}^n$ são convexos então o produto cartesiano $X \times Y \subset \mathbb{R}^{m+n}$ é convexo. Um exemplo de conjunto não-convexo é $X = \mathbb{R}^n - \{0\}$. Tem-se $e_1 \in X$, $-e_1 \in X$ mas $0 \in [-e_1, e_1]$.

Teorema 2. *Toda bola $B \subset \mathbb{R}^n$ é convexa.*

Demonstração: Seja $B = B(a; r)$ a bola aberta de centro a e raio $r > 0$. Se $x, y \in B$ então $|x - a| < r$ e $|y - a| < r$. Para qualquer $t \in [0, 1]$, temos:

$$\begin{aligned} |(1-t)x + ty - a| &= |(1-t)(x-a) + t(y-a)| \leq \\ &\leq (1-t)|x-a| + t|y-a| < r. \end{aligned}$$

Uma demonstração inteiramente análoga se faz no caso de uma bola fechada $B[a; r]$.

Note que a demonstração acima usa apenas as propriedades N1, N1 e N3 da norma, que pode ser arbitrária.

Um subconjunto $X \subset \mathbb{R}^n$ diz-se *limitado* quando existe um número real $c > 0$ tal que $|x| \leq c$ para todo $x \in X$. Isto equivale a dizer que X está contido na bola fechada de centro na origem e raio c.

Se existir alguma bola $B[a; r]$, de centro arbitrário, contendo X então, para todo $x \in X$, tem-se $|x - a| \leq r$. Pondo $c = r + |a|$ temos então

$$x \in X \Rightarrow |x| = |x - a + a| \leq |x - a| + |a| \leq r + |a| = c,$$

logo X é limitado. Assim, um conjunto $X \subset \mathbb{R}^n$ é limitado se, e somente se, está contido em alguma bola (cujo centro não é necessariamente a origem).

Em relação às três normas usuais do espaço $\mathbb{R}^n$, as desigualdades $|x|_M \leq |x| \leq |x|_S \leq n \cdot |x|_M$ mostram que um conjunto $X \subset \mathbb{R}^n$ é

limitado em relação a uma dessas normas se, e somente se, é limitado em relação a qualquer das outras duas.

Para cada $i = 1, 2, \ldots, n$, a i-ésima projeção $\pi_i \colon \mathbb{R}^n \to \mathbb{R}$ foi definida por $\pi_i(x) = x_i = i$-ésima coordenada de x.

Teorema 3. *Um conjunto $X \subset \mathbb{R}^n$ é limitado se, e somente se, suas projeções $X_1 = \pi_1(x), \ldots, X_n = \pi_n(X)$ são conjuntos limitados em $\mathbb{R}$.*

Demonstração: Começamos com uma observação preliminar: dados os conjuntos $C_1, \ldots, C_n \subset \mathbb{R}$ e $X \subset \mathbb{R}^n$, tem-se $X \subset C_1 \times \cdots \times C_n \Leftrightarrow \pi_1(X) \subset C_1, \ldots, \pi_n(X) \subset C_n$. Para demonstrar o teorema, podemos escolher em $\mathbb{R}^n$ qualquer das 3 normas usuais. Tomemos a norma do máximo. Então $X \subset \mathbb{R}^n$ é limitado $\Leftrightarrow$ para algum $c > 0$, tem-se $X \subset B[0; c] = [-c, c] \times \cdots \times [-c, c]$ (n fatores) $\Leftrightarrow \pi_1(X) \subset [-c, c], \ldots, \pi_n(X) \subset [-c, c] \Leftrightarrow \pi_1(X), \ldots, \pi_n(X)$ são limitados em $\mathbb{R}$.

5 Seqüências no espaço euclidiano

Uma seqüência em $\mathbb{R}^n$ é uma aplicação $x \colon \mathbb{N} \to \mathbb{R}^n$, definida no conjunto $\mathbb{N}$ dos números naturais. O valor que essa aplicação assume no número k é indicado com x_k e chama-se o *k-ésimo termo* da seqüência. Usaremos as notações (x_k), $(x_1, x_2, \ldots, x_k, \ldots)$ ou $(x_k)_{k\in\mathbb{N}}$ para indicar a seqüência cujo k-ésimo termo é $x_k \in \mathbb{R}^n$.

Uma *subseqüência* de (x_k) é a restrição da seqüência a um subconjunto infinito $\mathbb{N}' = \{k_1 < k_2 < \cdots < k_i < \cdots\} \subset \mathbb{N}$. A subseqüência é indicada pelas notações $(x_k)_{k\in\mathbb{N}'}$ ou $(x_{k_i})_{i\in\mathbb{N}}$ ou $(x_{k_1}, x_{k_2}, \ldots, x_{k_i}, \ldots)$.

Diz-se que a seqüência (x_k) é *limitada* quando o conjunto dos seus termos é limitado em $\mathbb{R}^n$, ou seja, quando existe um número real $c > 0$ tal que $|x_k| \leq c$ para todo $k \in \mathbb{N}$.

Uma seqüência (x_k) em $\mathbb{R}^n$ equivale a n seqüências de números reais. Com efeito, para cada $k \in \mathbb{N}$ temos $x_k = (x_{k1}, x_{k2}, \ldots, x_{kn})$ onde $x_{ki} = \pi_i(x_k) = i$-ésima coordenada de x_k ($i = 1, 2, \ldots, n$). As n seqüências $(x_{ki})_{k\in\mathbb{N}}$ ($i = 1, \ldots, n$) são chamadas as *seqüências das coordenadas* de (x_k). Assim, por exemplo, no plano $\mathbb{R}^2$, uma seqüência de pontos $z_k = (x_k, y_k)$ é o mesmo que um par de seqüências (x_k), (y_k) de números reais.

Segue-se imediatamente do Teorema 3 que uma seqüência (x_k) em $\mathbb{R}^n$ é limitada se, e somente se, cada uma de suas seqüências de coordenadas $(x_{ki})_{k\in\mathbb{N}}$ ($1 \leq i \leq n$) é limitada em $\mathbb{R}$.

Diz-se que o ponto $a \in \mathbb{R}^n$ é o *limite* da seqüência de pontos $x_k \in \mathbb{R}^n$ quando, para todo $\varepsilon > 0$ dado, é possível obter $k_0 \in \mathbb{N}$ tal que $k > k_0 \Rightarrow |x_k - a| < \varepsilon$. Neste caso, diz-se também que (x_k) *converge* para a ou *tende* para a, escreve-se $\lim x_k = a$, $\lim\limits_{k\to\infty} x_k = a$, $\lim\limits_{k\in\mathbb{N}} x_k = a$, ou simplesmente $x_k \to a$.

Quando existe o limite $a = \lim x_k$, diz-se que a seqüência (x_k) é *convergente*. Caso contrário, diz-se que (x_k) é *divergente*.

Por exemplo, uma seqüência constante $(a, a, \dots, a, \dots)$ é obviamente convergente e seu limite é a. Por outro lado, se $a \neq b$ então $(a, b, a, b, \dots)$ é uma seqüência divergente.

Tem-se $\lim x_k = a \Leftrightarrow \lim |x_k - a| = 0$. Isto reduz a convergência em $\mathbb{R}^n$ à convergência de números reais ≥ 0.

(5.1) Em termos de bolas, tem-se $\lim x_k = a$ se, e somente se, qualquer bola aberta de centro a contém todos os termos x_k salvo, possivelmente para um número finito de índices k. (Com efeito, se $\varepsilon > 0$ é o raio da bola e k_0 é o número natural que corresponde a ε na definição de limite, fora da bola $B(a;\varepsilon)$ só poderão estar, no máximo, alguns dos termos $x_1, \dots, x_{k_0}$.)

(5.2) Resulta da observação acima que toda seqüência convergente é limitada. De fato, se $\lim x_k = a$ então fora da bola aberta de centro a e raio 1 existem no máximo os termos $x_1, \dots, x_{k_0}$ da seqüência. Se r é o maior dos números 1, $|x_1 - a|, \dots, |x_{k_0} - a|$, vemos que todos os termos da seqüência estão contidos na bola $B[a;r]$.

A recíproca é falsa: se $a \neq b$, a seqüência $(a, b, a, b, \dots)$ é divergente e limitada.

(5.3) Segue-se também da caracterização do limite por meio de bolas que se $\lim x_k = a$ então toda subseqüência de (x_k) tem ainda limite igual a a. Ou seja: toda subseqüência de uma seqüência convergente é ainda convergente e tem o mesmo limite.

(5.4) Outro fato elementar, porém essencial, é que o limite de uma seqüência convergente é único. Ou seja, se $\lim x_k = a$ e $\lim x_k = b$, então $a = b$. Com efeito, para todo $k \in \mathbb{N}$ temos:

$$0 \leq |a - b| \leq |x_k - a| + |x_k - b|.$$

Logo $\lim |x_k - a| = \lim |x_k - b| = 0 \Rightarrow a = b$.

Em particular, se $\lim x_k = a$ e uma subseqüência de (x_k) converge para o ponto $b \in \mathbb{R}^n$, então $a = b$.

A definição de limite de uma seqüência em $\mathbb{R}^n$ faz uso de uma norma. As desigualdades que relacionam as 3 normas usuais do espaço euclidiano nos dão, entretanto: $|x_k-a|_M \leq |x_k-a| \leq |x_k-a|_S \leq n|x-a|_M$. Segue-se daí que $\lim |x_k - a|_M = 0 \Leftrightarrow \lim |x_k - a| = 0 \Leftrightarrow \lim |x_k - a|_S = 0$. Portanto, a afirmação $\lim x_k = a$ independe de qual das 3 normas usuais estamos considerando. (No final deste parágrafo, mostraremos que a noção de limite de uma seqüência em $\mathbb{R}^n$ permanece a mesma *seja qual for a norma* que considerarmos.)

Na demonstração do teorema abaixo, usamos em $\mathbb{R}^n$ a norma do máximo, por conveniência. Resulta do que acabamos de dizer que o fato nele enunciado é válido seja qual for a norma, dentre as 3 usuais, que seja tomada na definição de limite.

Teorema 4. *Uma seqüência (x_k) em $\mathbb{R}^n$ converge para o ponto $a = (a_1, \ldots, a_n)$ se, e somente se, para cada $i = 1, 2, \ldots, n$, tem-se $\lim\limits_{k\to\infty} x_{ki} = a_i$, ou seja, cada coordenada de x_k converge para a coordenada correspondente de a.*

Demonstração: Como $|x_{ki} - a_i| \leq |x_k - a|$, vemos que $\lim\limits_{k\to\infty} x_k = a \Rightarrow \lim\limits_{k\to\infty} x_{ki} = a_i$ para cada $i = 1, \ldots, n$. Reciprocamente, $\lim\limits_{k\to\infty} x_{ki} = a_i$ para todo $i = 1, 2, \ldots, n$ então, dado $\varepsilon > 0$, existem números naturais $k_1, \ldots, k_n$ tais que $k > k_i \Rightarrow |x_{ki} - a_i| < \varepsilon$. Seja $k_0 = \max\{k_1, \ldots, k_n\}$. Então $k > k_0 \Rightarrow |x_k - a| = \max\limits_i |x_{ki} - a_i| < \varepsilon$. Logo $\lim x_k = a$.

Corolário. *Dadas as seqüências convergentes de termos $x_k, y_k \in \mathbb{R}^n$ e $\alpha_k \in \mathbb{R}$, sejam $\lim x_k = a$, $\lim y_k = b$ e $\lim \alpha_k = \alpha$. Então:*

1) $\lim(x_k + y_k) = a + b$;
2) $\lim \alpha_k \cdot x_k = \alpha \cdot a$;
3) $\lim \langle x_k, y_k\rangle = \langle a, b\rangle$;
4) $\lim |x_k| = |a|$.

Com efeito, utilizando os fatos conhecidos sobre limites de somas e de produtos de números reais, vemos que, para cada $i = 1, \ldots, n$, valem $\lim\limits_{k\to\infty} (x_{ki} - y_{ki}) = a_i + b_i$ e $\lim\limits_{k\to\infty} \alpha_k x_{ki} = \alpha \cdot a_i$. Em vista do Teorema 4, as igualdades 1) e 2) ficam demonstradas. Analogamente, $\lim\limits_{k\to\infty} \langle x_k, y_k\rangle = \lim\limits_{k\to\infty} (x_{k1}y_{k1} + \cdots + x_{kn}y_{kn}) = a_1b_1 + \cdots + a_nb_n = \langle a, b\rangle$. Isto prova 3). Finalmente, $\lim |x_k| = \lim \sqrt{\langle x_k, x_k\rangle} = \sqrt{\langle x_k, x_k\rangle} = \sqrt{\langle a, a\rangle} = |a|$, o que prova 4). Também poderíamos provar 4) observando que $|\,|x_k| - |a|\,| \leq |x_k - a|$. Esta maneira tem a vantagem de valer para qualquer norma, euclidiana ou não.

Teorema 5. (Bolzano-Weierstrass). *Toda seqüência limitada em* $\mathbb{R}^n$ *possui uma subseqüência convergente.*

Demonstração: Sabemos que o Teorema de Bolzano-Weierstrass é válido na reta: toda seqüência limitada de números reais possui uma subseqüência convergente. (Vol. 1, pag. 123.) Dada a seqüência limitada (x_k) em $\mathbb{R}^n$, as primeiras coordenadas dos seus termos formam uma seqüência limitada $(x_{k1})_{k\in\mathbb{N}}$ de números reais, a qual possui uma subseqüência convergente. Isto é, existem um subconjunto infinito $\mathbb{N}_1 \subset \mathbb{N}$ e um número real a_1 tais que $\lim\limits_{k\in\mathbb{N}_1} x_{k1} = a_1$. Por sua vez, a seqüência limitada $(x_{k2})_{k\in\mathbb{N}_1}$ de números reais, possui uma subseqüência convergente; podemos obter um subconjunto infinito $\mathbb{N}_2 \subset \mathbb{N}_1$ e $a_2 \in \mathbb{R}$ tais que $\lim\limits_{k\in\mathbb{N}_2} x_{k2} = a_2$. E assim por diante, até encontrarmos conjuntos infinitos $\mathbb{N} \supset \mathbb{N}_1 \supset \mathbb{N}_2 \supset \cdots \supset \mathbb{N}_n$ e números reais $a_1, \ldots, a_n$ tais que $\lim\limits_{k\in\mathbb{N}_i} x_{ki} = a_i$ para $i = 1, 2, \ldots, n$. Então pomos $a = (a_1, \ldots, a_n)$ e vemos, pelo Teorema 4, que $\lim\limits_{k\in\mathbb{N}_n} x_k = a$, o que conclui a demonstração.

Diz-se que um ponto $a \in \mathbb{R}^n$ é *valor de aderência* de uma seqüência de pontos $x_k \in \mathbb{R}^n$ quando alguma subseqüência de (x_k) converge para a. O Teorema 5 diz que o conjunto dos valores de aderência de uma seqüência limitada em $\mathbb{R}^n$ nunca é vazio. Se uma seqüência (x_k) em $\mathbb{R}^n$ (necessariamente ilimitada) não possui valores de aderência, escreve-se $\lim x_k = \infty$. Isto significa que (x_k) não possui subseqüência limitada, ou seja, que para todo número real A, dado arbitrariamente, existe $k_0 \in \mathbb{N}$ tal que $k > k_0 \Rightarrow |x_k| > A$.

Do mesmo modo como para seqüência de números reais (Vol. 1, pág. 121) se mostra que o ponto $a \in \mathbb{R}^n$ é valor de aderência da seqüência (x_k) se, e somente se, toda bola de centro a contém termos x_k com índices arbitrariamente grandes. Mais precisamente: dados $\varepsilon > 0$ e $k_0 \in \mathbb{N}$, existe $k > k_0$ tal que $|x_k - a| < \varepsilon$.

Uma seqüência convergente possui um único valor de aderência. A recíproca não vale: a seqüência de números reais $(1, 2, 1, 3, 1, 4, 1, 5, \ldots)$ possui o único valor de aderência 1 mas não converge. Temos, entretanto o

Teorema 6. *Uma seqüência limitada em* $\mathbb{R}^n$ *é convergente se, e somente se, possui um único valor de aderência.*

Demonstração: Seja a um valor de aderência da seqüência limitada (x_k). Se não for $a = \lim x_k$ então existe $\varepsilon > 0$ tal que o conjunto

$\mathbb{N}' = \{k \in \mathbb{N}; x_k \notin B(a;\varepsilon)\}$ é infinito. A subseqüência $(x_k)_{k\in\mathbb{N}'}$ é limitada e seus termos satisfazem todos a condição $|x_k - a| \geq \varepsilon$. Pelo Teorema 5, ela possui uma subseqüência que converge para um ponto $b \in \mathbb{R}^n$. Ora, de $|x_k - a| \geq \varepsilon$ para todo $k \in \mathbb{N}'$, concluímos por passagem ao limite que $|b - a| \geq \varepsilon$. Então $b \neq a$ e a seqüência (x_k) tem pelo menos dois valores de aderência distintos. Isto demonstra metade do Teorema 6. A recíproca é imediata.

Estenderemos agora ao espaço euclidiano o Critério de Cauchy para convergência de seqüências de números reais.

Uma seqüência (x_k) em $\mathbb{R}^n$ diz-se uma *seqüência de Cauchy* quando para todo $\varepsilon > 0$ existe $k_0 \in \mathbb{N}$ tal que $k, r > k_0 \Rightarrow |x_k - x_r| < \varepsilon$.

Usando em $\mathbb{R}^n$ a norma do máximo, temos

$$|x_k - x_r| = \max\{|x_{k1} - x_{r1}|, \ldots, |x_{kn} - x_{rn}|\},$$

logo (x_k) é uma seqüência de Cauchy em $\mathbb{R}^n$ se, e somente se, para cada $i = 1, \ldots, n$, a seqüência $(x_{1i}, x_{2i}, \ldots, x_{ki}, \ldots)$ das suas i-ésimas coordenadas é uma seqüência de Cauchy de números reais. Isto conduz ao

Teorema 7. *Uma seqüência (x_k) em $\mathbb{R}^n$ é de Cauchy se, e somente se, é convergente.*

Demonstração: Se (x_k) é uma seqüência de Cauchy então, para cada $i = 1, 2, \ldots, n$, suas i-ésimas coordenadas formam uma seqüência de Cauchy de números reais $(x_{ki})_{k\in\mathbb{N}}$, a qual possui o limite $a_i = \lim_{k\to\infty} x_{ki}$. (Vol. 1, pág. 127.) Seja $a = (a_1, \ldots, a_n)$. Resulta do Teorema 4 acima que $a = \lim x_k$. Logo toda seqüência de Cauchy em $\mathbb{R}^n$ é convergente. A recíproca é imediata.

Generalizando as desigualdades que relacionam as 3 normas usuais do espaço euclidiano, diremos que duas normas arbitrárias $|\ |$ e $||\ ||$ em $\mathbb{R}^n$ são *equivalentes* quando existirem constantes $a > 0$ e $b > 0$ tais que

$$|x| \leq a||x|| \ \text{ e } \ ||x|| \leq b|x| \text{ para todo } x \in \mathbb{R}^n.$$

Se, para todo $x \in \mathbb{R}^n$ e todo $r > 0$, indicarmos com as notações $B(x,r)$ e $B((x;r))$ respectivamente a bola aberta de centro x e raio r segundo as normas $|\ |$ e $||\ ||$, as desigualdades acima significam que $B((x;r)) \subset B(x;ar)$ e $B(x;r) \subset B((x;br))$.

A equivalência entre normas é uma relação reflexiva, simétrica e transitiva. As 3 normas usuais do espaço euclidiano são equivalentes umas às outras. É claro que se $|\ |$ e $||\ ||$ são equivalentes, então $\lim |x_k - a| = 0 \Leftrightarrow \lim ||x_k - a|| = 0$, isto é, normas equivalentes dão origem à mesma noção de limite no espaço $\mathbb{R}^n$. Finalmente, se duas normas em $\mathbb{R}^n$ são equivalentes, um conjunto $X \subset \mathbb{R}^n$ é limitado em relação a uma delas se, e somente se, é limitado em relação à outra.

Demonstraremos agora o principal resultado a respeito de equivalência entre normas.

Teorema 8. *Duas normas quaisquer no espaço $\mathbb{R}^n$ são equivalentes.*

Demonstração: Seja $|x| = \sum_{i=1}^{n} |x_i|$ a norma da soma. Por transitividade, basta demonstrar que uma norma arbitrária $||x||$ em $\mathbb{R}^n$ é equivalente a esta. Em primeiro lugar, seja $b = \max\{||e_1||, \ldots, ||e_n||\}$. Então, para qualquer $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ temos

$$||x|| = ||x_1e_1 + \cdots + x_ne_n|| \leq |x_1|\,||e_1|| + \cdots + |x_n|\,||e_n|| \leq b \cdot |x|.$$

Resta provar que existe $a > 0$ tal que $|x| \leq a||x||$ para todo $x \in \mathbb{R}^n$. Suponha, por absurdo, que não seja assim. Então, para cada $k \in \mathbb{N}$, podemos achar $x_k \in \mathbb{R}^n$ tal que $|x_k| > k \cdot ||x_k||$. Ponhamos $u_k = x_k/|x_k|$. Isto nos dá $||u_k|| = ||x_k||/|x_k| < 1/k$ e $|u_k| = 1$ para todo k. A seqüência (u_k) é, portanto, limitada em relação à norma da soma. Pelo Teorema de Bolzano-Weierstrass, ela possui uma subseqüência (u_{k_j}) que converge para um ponto $u \in \mathbb{R}^n$. Por um lado, temos $|u| = \lim_{j\to\infty} |u_{k_j}| = 1$, donde $u \neq 0$. Por outro lado, para todo $j \in \mathbb{N}$ temos

$$||u|| \leq ||u_{k_j} - u|| + ||u_{k_j}|| \leq b|u_{k_j} - u| + 1/k_j\,.$$

Como as duas últimas parcelas acima tendem para zero quando $j \to \infty$, concluimos que $||u|| = 0$, donde $u = 0$. Esta contradição demonstra o teorema.

Praticamente todos os conceitos introduzidos e por introduzir, bem como resultados demonstrados ou a demonstrar neste livro, têm sua extensão e sua validez inalteradas se substituirmos a norma que estamos usando por outra equivalente. O Teorema 8 então diz que para esses fatos, podemos usar *qualquer* norma em $\mathbb{R}^n$. Continuaremos porém admitindo que, salvo menção explícita em contrário, a norma adotada em $\mathbb{R}^n$ é a euclidiana. Isto tem a vantagem de deixar um produto interno

à nossa disposição. Quando houver conveniência, lançaremos mão de outra norma, que será quase sempre uma das que chamamos "usuais".

Exemplo. Como aplicação do teorema acima, mostraremos que uma seqüência de polinômios $p_k(t) = a_{k0} + a_{k1}t + \cdots + a_{kn}t^n$, todos de grau $\leq n$, converge para o polinômio $p(t) = a_0 + a_1t + \cdots + a_nt^n$, uniformemente no intervalo não degenerado $[\alpha, \beta]$, se, e somente se, para cada $i = 0, 1, \ldots, n$, a seqüência $(a_{ki})_{k\in\mathbb{N}}$ dos coeficientes de t^i nos polinômios p_k converge para o coeficiente a_i de t^i no polinômio p. Com efeito, existe um isomorfismo entre $\mathbb{R}^{n+1}$ e o espaço vetorial dos polinômios reais de grau $\leq n$, o qual faz corresponder a cada ponto $x = (a_0, \ldots, a_n) \in \mathbb{R}^{n+1}$ o polinômio $p_x(t) = a_0 + a_1t + \cdots + a_nt^n$. Vê-se facilmente que $||x|| = \sup\{|p_x(t)|; t \in [\alpha, \beta]\}$ define uma norma em $\mathbb{R}^{n+1}$. Em relação a esta norma, $x_k \to a$ em $\mathbb{R}^{n+1}$ significa que $p_{x_k} \to p_a$ uniformemente no intervalo $[\alpha, \beta]$. Como duas normas quaisquer em $\mathbb{R}^{n+1}$ são equivalentes, pondo $x_k = (a_{k0}, a_{k1}, \ldots, a_{kn})$ e $a = (a_0, a_1, \ldots, a_n)$ vem: $\lim\limits_{k\to\infty} a_{ki} = a_i$ para todo $i \Leftrightarrow \lim ||x_k - a|| = 0 \Leftrightarrow \lim p_k = p$ uniformemente em $[\alpha, \beta]$.

6 Pontos de acumulação

Seja $X \subset \mathbb{R}^n$. Um ponto $a \in \mathbb{R}^n$ chama-se *ponto de acumulação* do conjunto X quando toda bola aberta de centro a contém algum ponto de X, diferente do ponto a. Noutros termos, para todo $\varepsilon > 0$, deve existir $x \in X$ tal que $0 < |x - a| < \varepsilon$.

Por exemplo, seja $X \subset \mathbb{R}^n$ a bola aberta de centro na origem e raio $r > 0$. Todo ponto $a \in \mathbb{R}^n$ com $|a| = r$ é ponto de acumulação de X. Com efeito, dado $\varepsilon > 0$, podemos, sem perda de generalidade, supor $\varepsilon < 2r$. Então o ponto $x = (1 - \varepsilon/2r)a$ pertence à bola X, é diferente de a e tem-se $|a - x| = \varepsilon/2 < \varepsilon$.

O conjunto dos pontos de acumulação de X será representado pela notação X'. X' chama-se o *conjunto derivado* de X.

Como vimos no exemplo acima, um ponto de acumulação do conjunto X pode não pertencer a X.

Teorema 9. *Dados $X \subset \mathbb{R}^n$ e $a \in \mathbb{R}^n$, as seguintes afirmações são equivalentes:*

1) *a é ponto de acumulação de X;*

2) *Existe uma seqüência de pontos* $x_k \in X$, *com* $\lim x_k = a$ *e* $x_k \neq a$ *para todo* $k \in \mathbb{N}$*;*

3) *Toda bola aberta de centro* a *contém uma infinidade de pontos de* X*.*

Demonstração: Supondo 1), para todo $k \in \mathbb{N}$, obtemos um ponto $x_k \in X$ tal que $0 < |x_k - a| < 1/k$, donde $x_k \neq a$ e $\lim x_k = a$. Logo 1) $\Rightarrow$ 2). Por outro lado, supondo 2) então para qualquer $k_0 \in \mathbb{N}$ o conjunto $\{x_k; k > k_0\}$ é infinito porque do contrário haveria um termo x_k que se repetiria infinitas vezes e isto forneceria uma subseqüência (constante) convergindo para o limite $x_k \neq a$. Então 2) $\Rightarrow$ 3). Finalmente, é óbvio que 3) implica 1).

Corolário. *Se* $X' \neq \varnothing$ *então* X *é infinito.*

Um conjunto infinito pode não possuir ponto de acumulação. Tal é o caso do conjunto $\mathbb{Z} \subset \mathbb{R}$ dos números inteiros. Vale entretanto o teorema abaixo, também chamado "*Teorema de Bolzano-Weierstrass*".

Teorema 10. *Se* $X \subset \mathbb{R}^n$ *é infinito e limitado então* $X' \neq \varnothing$*.*

Demonstração: Sendo infinito, X contém um subconjunto enumerável $\{x_1, x_2, \ldots, x_k, \ldots\}$. Obtemos assim uma seqüência (x_k), a qual é limitada, logo, pelo Teorema 5 (Bolzano-Weierstrass), admite uma subseqüência que converge para um ponto $a \in \mathbb{R}^n$. Como os termos x_k são 2 a 2 distintos, no máximo um deles é igual a a. Eliminando-o, se necessário, obtemos uma seqüência de pontos de X, todos diferentes de a, com limite a. Pelo Teorema 9, a é ponto de acumulação de X.

Se $a \in X$ não é ponto de acumulação de X, diz-se que a é um *ponto isolado* de X. Para que isso aconteça, é necessário e suficiente que exista $\varepsilon > 0$ com $B(a;\varepsilon) \cap X = \{a\}$.

Quando todo ponto $a \in X$ é isolado, dizemos que X é um conjunto *discreto*.

7 Aplicações contínuas

Seja $f\colon X \to \mathbb{R}^n$ uma aplicação definida no conjunto $X \subset \mathbb{R}^m$. Diz-se que f é *contínua* no ponto $a \in X$ quando, para qualquer $\varepsilon > 0$ dado, se pode obter $\delta > 0$ tal que todo ponto $x \in X$ cuja distância ao ponto a seja menor do que δ é transformado por f num ponto $f(x)$ que dista de

$f(a)$ menos que ε. Em linguagem simbólica

$$\forall \varepsilon > 0 \, \exists \delta > 0; \ x \in X, \ |x - a| < \delta \Rightarrow |f(x) - f(a)| < \varepsilon.$$

Em termos de bolas, a continuidade de f no ponto a se exprime assim: para toda bola aberta B' de centro $f(a)$ em $\mathbb{R}^n$ existe uma bola aberta B de centro a em $\mathbb{R}^m$ tal que $f(B \cap X) \subset B'$.

Se $f\colon X \to \mathbb{R}^n$ é contínua em todos os pontos do conjunto X, diz-se simplesmente que f é uma aplicação *contínua*.

A continuidade é um fenômeno local: se cada ponto $a \in X$ é centro de uma bola B tal que a restrição $f \mid (B \cap X)$ é contínua, então $f\colon X \to \mathbb{R}^n$ é contínua.

Embora a definição de continuidade de uma aplicação $f\colon X \to \mathbb{R}^n$ ($X \subset \mathbb{R}^m$) faça uso de uma norma em $\mathbb{R}^m$ e outra em $\mathbb{R}^n$ (ambas indicadas acima com a mesma notação), segue-se da definição de normas equivalentes e do Teorema 8 que a continuidade (ou descontinuidade) de f num ponto persiste se alterarmos uma dessas normas ou ambas.

Como se vê facilmente, se $f\colon X \to \mathbb{R}^n$ é contínua então, para $Y \subset X$, a restrição $f \mid Y$ é uma aplicação contínua.

Um caso trivial de continuidade é o seguinte: se a é um ponto isolado do conjunto X, então *toda* aplicação $f\colon X \to \mathbb{R}^n$ é necessariamente contínua no ponto a. Com efeito, existe $\delta > 0$ tal que $B(a;\delta) \cap X = \{a\}$. Assim, para qualquer $\varepsilon > 0$ dado, tomamos este valor de δ e temos $x \in X, \ \ |x - a| < \delta \Rightarrow x = a \Rightarrow |f(x) - f(a)| = 0 < \varepsilon$.

Vejamos agora outros exemplos de aplicações contínuas.

Dado $X \subset \mathbb{R}^m$, uma aplicação $f\colon X \to \mathbb{R}^n$ diz-se *Lipschitziana* quando existe $k > 0$ (constante de Lipschitz de f) tal que, para quaisquer $x, y \in X$, tem-se $|f(x) - f(y)| \leq k|x - y|$. Toda aplicação Lipschitziana é contínua: dado $\varepsilon > 0$, basta tomar $\delta = \varepsilon/k$. Então $|x - a| < \delta \Rightarrow |f(x) - f(a)| \leq k|x - a| < k \cdot \varepsilon/k = \varepsilon$. Ser ou não Lipschitziana independe das normas.

(7.1) Toda transformação linear $A\colon \mathbb{R}^m \to \mathbb{R}^n$ é Lipschitziana. Com efeito, seja $c = \max\{|A \cdot e_1|, \cdots, |A \cdot e_m|\}$. Então, para todo $x = (x_1, \ldots, x_m) \in \mathbb{R}^m$ temos

$$|A \cdot x| = |A \cdot (\Sigma\, x_i e_i)| = |\Sigma\, x_i A \cdot e_i| \leq \Sigma |x_i| \cdot |A \cdot e_i| \leq c \cdot \Sigma |x_i|.$$

Tomando em $\mathbb{R}^m$ a norma da soma, temos $|A \cdot x| \leq c|x|$ para todo $x \in \mathbb{R}^m$. Então, para $x, y \in \mathbb{R}^m$ arbitrários, vale:

$$|A \cdot x - A \cdot y| = |A \cdot (x - y)| \leq c|x - y|.$$

Assim, A cumpre a condição de Lipschitz, com uma constante igual à maior das normas dos vetores-coluna de sua matriz. Em particular, A é contínua.

(7.2) Seja agora $\varphi\colon \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^p$ uma aplicação bilinear. Salvo no caso trivial em que é identicamente nula, φ não é Lipschitziana mas, conforme mostraremos agora, goza dessa propriedade em cada parte limitada de $\mathbb{R}^m \times \mathbb{R}^n = \mathbb{R}^{m+n}$. Consideremos neste espaço a norma da soma e chamemos de c o maior dos números $|\varphi(e_i, e_j)|$, $1 \le i \le m$, $1 \le j \le n$. Para quaisquer $x \in \mathbb{R}^m$, $y \in \mathbb{R}^n$, temos $x = \sum_i x_i e_i$, $y = \sum_j y_j e_j$, $|x|\,|y| = \sum_{i,j} |x_i|\,|y_j|$ e $\varphi(x, y) = \sum_{i,j} x_i y_j\, \varphi(e_i, e_j)$. Logo

$$|\varphi(x,y)| \le \sum_{i,j} |x_i|\,|y_i|\,|\varphi(e_i,e_j)| \le c \cdot \sum_{i,j} |x_i|\,|y_j| = c \cdot |x|\,|y|.$$

Então, dados arbitrariamente $z = (x, y)$ e $z' = (x', y')$ em $\mathbb{R}^m \times \mathbb{R}^n$, vale:

$$\begin{aligned} &|\varphi(z) - \varphi(z')| = |\varphi(x, y - y') + \varphi(x - x', y')| \le \\ &\le |\varphi(x, y - y')| + |\varphi(x - x', y')| \le c(|x|\,|y - y'| + |y'|\,|x - x'|). \end{aligned}$$

Se z e z' pertencem ambos à bola $B[0, r]$ em $\mathbb{R}^{m+n}$, temos, em particular, $|x| \le r$ e $|y'| \le r$. Assim:

$$|\varphi(z) - \varphi(z')| \le c \cdot r \cdot (|x - x'| + |y - y'|) = c \cdot r \cdot |z - z'|.$$

Portanto φ cumpre uma condição de Lipschitz (com constante $c \cdot r$) em cada bola $B[0, r]$ do espaço $\mathbb{R}^m \times \mathbb{R}^n = \mathbb{R}^{m+n}$. Em particular, toda aplicação bilinear é contínua. Casos especiais de aplicações bilineares são: a multiplicação de números reais, $\varphi\colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, $\varphi(x, y) = xy$; a multiplicação de um escalar por um vetor, $\varphi\colon \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$, $\varphi(\alpha, x) = \alpha \cdot x$; o produto interno $\varphi\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $\varphi(x, y) = \Sigma\, x_i y_i$; a multiplicação de matrizes $\varphi\colon \mathbb{R}^{mn} \times \mathbb{R}^{np} \to \mathbb{R}^{mp}$, $\varphi(X, Y) = X \cdot Y$, e a "avaliação" $\varphi\colon \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n) \times \mathbb{R}^m \to \mathbb{R}^n$, $\varphi(A, x) = A(x)$. Segue-se que todas essas aplicações são contínuas. (No caso da última, identifica-se $\mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$ com $\mathbb{R}^{mn}$.)

Um caso particular de aplicações Lipschitzianas é o das *imersões isométricas*, que são aplicações $f\colon X \to \mathbb{R}^n$ $(X \subset \mathbb{R}^m)$ que preservam as distâncias, isto é, cumprem $|f(x) - f(y)| = |x - y|$ para quaisquer $x, y \in X$. Uma imersão isométrica $f\colon X \to \mathbb{R}^n$ é sempre injetiva, pois $f(x) = f(y) \Rightarrow |x - y| = |f(x) - f(y)| = 0$, donde $x = y$. Exemplo

de imersão isométrica: para $m < n$ seja $f\colon \mathbb{R}^m \to \mathbb{R}^n$ dada por $f(x) = f(x_1, \ldots, x_m) = (x_1, \ldots, x_m, 0, \ldots, 0)$.

Uma imersão isométrica $f\colon X \to \mathbb{R}^n$, com $f(X) = Y$, chama-se uma *isometria* de X sobre Y. Sua inversa $f^{-1}\colon Y \to \mathbb{R}^m$ é uma isometria de Y sobre X.

Exemplos particularmente simples de isometrias são as translações $T_a\colon \mathbb{R}^m \to \mathbb{R}^m$, definidas com auxílio de um vetor fixo $a \in \mathbb{R}^m$, pondo-se $T_a(x) = x + a$. Evidentemente cada translação T_a é bijetiva, pois $T_{-a} = (T_a)^{-1}$. As translações não são aplicações lineares, salvo $T_0 =$ identidade. Uma transformação linear $A\colon \mathbb{R}^m \to \mathbb{R}^m$ é uma isometria se, e somente se, é *ortogonal*, ou seja $\langle Ax, Ay\rangle = \langle x, y\rangle$ para $x, y \in \mathbb{R}^m$ arbitrários. Uma transformação ortogonal A se caracteriza pelo fato de que os transformados $(A \cdot e_1, \ldots, A \cdot e_m)$ dos vetores da base canônica de $\mathbb{R}^m$ formam uma base ortonormal de $\mathbb{R}^m$. Isto equivale a dizer que as colunas da matriz de A são duas a duas ortogonais, de comprimento 1, ou seja que o produto $A^* \cdot A$, da transposta A^* por A, é igual à matriz identidade.

Uma *contração fraca* $f\colon X \to \mathbb{R}^n$ é uma aplicação Lipschitziana com $k = 1$, ou seja, com $|f(x) - f(y)| \le |x - y|$ para quaisquer $x, y \in X \subset \mathbb{R}^m$.

Exemplos de contrações fracas são:

a) a *soma de vetores* $s\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$, definida por $s(x, y) = x + y$. De fato, tomando em $\mathbb{R}^n \times \mathbb{R}^n = \mathbb{R}^{2n}$ a norma da soma, vem

$$|s(x, y) - s(x', y')| = |(x + y) - (x' + y')| = |(x - x') + (y - y')| \le$$

$$\le |x - x'| + |y - y'| = |(x, y) - (x', y')|.$$

b) As *projeções* $\pi_i\colon \mathbb{R}^m \to \mathbb{R}$, definidas por $\pi_i(x) = x_i$, onde $x = (x_1, \ldots, x_m)$. Tem-se $|\pi_i(x) - \pi_i(y)| = |x_i - y_i| \le |x - y|$, podendo-se tomar em $\mathbb{R}^m$ qualquer das 3 normas usuais.

c) A norma $|\ |\colon \mathbb{R}^m \to \mathbb{R}$. Com efeito, para quaisquer $x, y \in \mathbb{R}^m$, tem-se $|\,|x| - |y|\,| \le |x - y|$. Daí resulta que a distância $d\colon \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$, definida por $d(x, y) = |x - y|$, também é uma contração fraca (desde que tomemos em $\mathbb{R}^m \times \mathbb{R}^m$ a norma da soma). Com efeito,

$$|d(x, y) - d(x', y')| = |\,|x - y| - |x' - y'|\,| \le |(x - y) - (x' - y')| =$$

$$= |x - x' + y' - y| \le |x - x'| + |y - y'| = |(x, y) - (x', y')|.$$

Em particular, as aplicações consideradas nos exemplos a), b) e c) são contínuas.

Se trocarmos a norma de $\mathbb{R}^m$ ou de $\mathbb{R}^n$, uma contração fraca continua sendo Lipschitziana (e portanto contínua) mas a constante muda e ela pode deixar de ser uma contração fraca.

Teorema 11. *A composta de duas aplicações contínuas é contínua. Mais precisamente, dados $X \subset \mathbb{R}^m$, $Y \subset \mathbb{R}^n$, $f\colon X \to \mathbb{R}^n$ contínua no ponto $a \in X$, com $f(X) \subset Y$, $g\colon Y \to \mathbb{R}^p$ contínua no ponto $b = f(a)$, então $g \circ f\colon X \to \mathbb{R}^p$ é contínua no ponto a.*

Demonstração: Dado arbitrariamente $\varepsilon > 0$ existe, em virtude da continuidade de g, um número $\eta > 0$ tal que $y \in Y$, $|y - f(a)| < \eta \Rightarrow |g(y) - gf(a))| < \varepsilon$. Por sua vez, a partir de η, a continuidade de f nos fornece $\delta > 0$ tal que $x \in X$, $|x - a| < \delta \Rightarrow |f(x) - f(a)| < \eta$. Segue-se que $x \in X$, $|x - a| < \delta \Rightarrow |g(f(x)) - g(f(a))| < \varepsilon$, logo $g \circ f$ é contínua no ponto a.

Seja $X \subset \mathbb{R}^m$. Dar uma aplicação $f\colon X \to \mathbb{R}^n$ é o mesmo que dar n funções reais $f_1, \ldots, f_n\colon X \to \mathbb{R}$, definidas por $f_i = \pi_i \circ f$, as quais são chamadas as *coordenadas* da aplicação f. Para todo $x \in X$, temos $f(x) = (f_1(x), \ldots, f_n(x))$. Para indicar que as f_i são as funções coordenadas de f, escreve-se $f = (f_1, \ldots, f_n)$.

Teorema 12. *Uma aplicação $f\colon X \to \mathbb{R}^n$, definida no conjunto $X \subset \mathbb{R}^m$, é contínua num ponto $a \in X$ se, e somente se, cada uma das suas funções coordenadas $f_i = \pi_i \circ f\colon X \to \mathbb{R}$ é contínua no ponto a.*

Demonstração: A continuidade de f implica a continuidade das f_i pelo teorema anterior. Reciprocamente, se cada $f_i\colon X \to \mathbb{R}^n$ é contínua no ponto $a \in X$, dado $\varepsilon > 0$ existem números $\delta_1 > 0, \ldots, \delta_n > 0$ tais que $|x - a| < \delta_i$, $x \in X \Rightarrow |f_i(x) - f_i(a)| < \varepsilon$. Tomemos em $\mathbb{R}^n$ a norma do máximo. Seja δ o menor dos δ_i. Então $|x - a| < \delta$, $x \in X \Rightarrow |f(x) - f(a)| = \max\{|f_i(x) - f_i(a)|; 1 \leq i \leq n\} < \varepsilon$. Logo f é contínua no ponto a.

Corolário. *Dadas $f\colon X \to \mathbb{R}^m$ e $g\colon X \to \mathbb{R}^n$, seja $(f,g)\colon X \to \mathbb{R}^m \times \mathbb{R}^n = \mathbb{R}^{m+n}$ definida por $(f,g)(x) = (f(x), g(x))$. Então (f,g) é contínua se, e somente se, f e g são ambas contínuas.*

Com efeito, se $f_1, \ldots, f_m$ são as funções coordenadas de f e $g_1, \ldots, g_n$ as de g, então as funções coordenadas de (f,g) são $f_1, \ldots, f_m$, $g_1, \ldots, g_n$.

Os Teoremas 11 e 12 são instrumentos de grande utilidade para estabelecer a continuidade de certas aplicações. Vejamos alguns exemplos para mostrar como se faz isso:

(7.3) Sejam $X \subset \mathbb{R}^m$ e $f, g\colon X \to \mathbb{R}^n$, $\alpha\colon X \to \mathbb{R}$ aplicações contínuas. Então são também contínuas as aplicações:

$$\begin{array}{ll} f+g\colon X \to \mathbb{R}^n, & (f+g)(x) = f(x)+g(x); \\ \alpha\cdot f\colon X \to \mathbb{R}^n, & (\alpha\cdot f)(x) = \alpha(x)\cdot f(x); \\ \langle f,g\rangle\colon X \to \mathbb{R}, & \langle f,g\rangle(x) = \langle f(x), g(x)\rangle; \\ 1/\alpha\colon X \to \mathbb{R}, & (1/\alpha)(x) = 1/\alpha(x) \text{ [definida se } 0 \notin \alpha(X)]. \end{array}$$

Com efeito, sejam $s\colon \mathbb{R}^n\times\mathbb{R}^n \to \mathbb{R}^n$ $\varphi\colon \mathbb{R}\times\mathbb{R}^n \to \mathbb{R}^n$, $\xi\colon \mathbb{R}^n\times\mathbb{R}^n \to \mathbb{R}$ e $\rho\colon \mathbb{R}-\{0\} \to \mathbb{R}$ dadas por $s(x,y) = x+y$, $\varphi(t,x) = t\cdot x$, $\xi(x,y) = \langle x,y\rangle$ e $\rho(t) = 1/t$. Então, usando a notação $(f,g)\colon X \to \mathbb{R}^n \times \mathbb{R}^n$ para indicar que $(f,g)(x) = (f(x), g(x))$, temos: $f+g = s\circ(f,g)$, $\alpha\cdot f = \varphi\circ(\alpha, f)$, $\langle f,g\rangle = \xi\circ(f,g)$ e $1/\alpha = \rho\circ\alpha$. Sabemos que s, φ, ξ e ρ (esta última, da Análise de uma variável) são contínuas. Usando os Teoremas 11 e 12, as afirmações ficam demonstradas.

Mais um exemplo: seja $f\colon \mathbb{R}^2 \to \mathbb{R}$ dada por $f(x,y) = (\operatorname{sen} x)\cdot e^{x^2+y^3}$. Mostrar que f é contínua. Ora, f é o produto das funções $g\colon \mathbb{R}^2 \to \mathbb{R}$ e $h\colon \mathbb{R}^2 \to \mathbb{R}$, onde $g(x,y) = \operatorname{sen} x$ e $h(x,y) = e^{x^2+y^3}$. Por um lado temos $g = \operatorname{sen}\circ\pi_1$, onde a projeção $\pi_1\colon \mathbb{R}^2 \to \mathbb{R}$ e a função $\operatorname{sen}\colon \mathbb{R} \to \mathbb{R}$ são contínuas. Por outro lado, $h = \exp\circ\varphi$, com $\exp(t) = e^t$ e $\varphi(x,y) = x^2+y^3$. Resta apenas mostrar que $\varphi\colon \mathbb{R}^2 \to \mathbb{R}$ é contínua. Mas φ é a soma das duas funções $(x,y) \mapsto x^2$ e $(x,y) \mapsto y^3$. As primeiras destas é a composta $\xi\circ\pi_1$, onde $\xi(x) = x^2$ e a segunda é a composta $\zeta\circ\pi_1$, onde $\xi(x) = x^2$ e a segunda é a composta $\zeta\circ\pi_2$, onde $\zeta(y) = y^3$. Logo φ é contínua, donde f também o é.

Na prática, verificações rotineiras como a que fizemos acima são substituídas pela simples observação de que, para obter $f(x,y)$ as coordenadas x e y submeteram-se a operações definidas por funções contínuas.

Dentro deste ponto de vista, a verificação de que uma aplicação linear $A\colon \mathbb{R}^m \to \mathbb{R}^n$ é contínua se faz assim: cada função coordenada $A_i\colon \mathbb{R}^m \to \mathbb{R}$ $(i = 1,\ldots,n)$ de A é um funcional linear. Ora, dado um funcional linear $f\colon \mathbb{R}^m \to \mathbb{R}$, tem-se $f(x) = a_1x_1 + \cdots + a_mx_m$, onde $a_i = f(e_i)$. Logo f é uma combinação linear, com coeficientes constantes, das projeções $x \mapsto x_i$ e portanto é contínua. Segue-se então que A é contínua.

Teorema 13. *Uma aplicação $f\colon X \to \mathbb{R}^n$, definida no subconjunto $X \subset \mathbb{R}^m$, é contínua no ponto $a \in X$ se, e somente se, para toda seqüência de pontos $x_k \in X$ com $\lim x_k = a$, tem-se $\lim\limits_{k\to\infty} f(x_k) = f(a)$.*

Demonstração: Sejam f contínua no ponto a e $\lim x_k = a$. Dado $\varepsilon > 0$, existe $\delta > 0$ tal que $x \in X$, $|x - a| < \delta \Rightarrow |f(x) - f(a)| < \varepsilon$. Como $\lim x_k = a$, existe $k_0 \in \mathbb{N}$ tal que $k > k_0 \Rightarrow |x_k - a| < \delta$. Segue-se que $k > k_0 \Rightarrow |f(x_k) - f(a)| < \varepsilon$. Logo $\lim f(x_k) = f(a)$. Para demonstrar a recíproca, suponhamos que f não seja contínua no ponto a. Então existe um $\varepsilon > 0$ tal que, para cada $k \in \mathbb{N}$ podemos obter $x_k \in X$ com $|x_k - a| < 1/k$ e $|f(x_k) - f(a)| \geq \varepsilon$. Então $\lim x_k = a$ sem que seja $\lim f(x_k) = f(a)$.

Em vista do Teorema 13, o corolário do Teorema 4 é apenas outra maneira de dizer que a soma, a multiplicação por um escalar e o produto interno em $\mathbb{R}^n$ são operações contínuas.

Uma aplicação $f\colon \mathbb{R}^m \to \mathbb{R}^n$ diz-se *contínua em relação à variável* x_i $(1 \leq i \leq m)$ quando, para cada $(a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_m)$ fixado, a "aplicação parcial" $t \mapsto f(a_1, \ldots, a_{i-1}, t, a_{i+1}, \ldots, a_m)$ é contínua.

Toda aplicação contínua $f\colon \mathbb{R}^m \to \mathbb{R}^n$ é separadamente contínua em relação a cada uma das suas variáveis, pois suas aplicações parciais são compostas de f com uma aplicação contínua do tipo $t \mapsto (a_1, \ldots, a_{i-1}, t, a_{i+1}, \ldots, a_m)$. Mas a recíproca é falsa: a função $f\colon \mathbb{R}^2 \to \mathbb{R}$, definida por $f(x, y) = xy/(x^2 + y^2)$ se $x^2 + y^2 \neq 0$ e $f(0, 0) = 0$, é contínua separadamente em relação a x e a y, pois $f(a, y) = ay/(a^2 + y^2)$ se $a \neq 0$ e $f(0, y) = 0$, enquanto $f(x, b) = bx/(x^2 + b^2)$ se $b \neq 0$ e $f(x, 0) = 0$. Mas f é descontínua na origem $(0, 0)$. Com efeito, se definirmos $g\colon \mathbb{R} \to \mathbb{R}^2$ por $g(t) = (at, at)$, então $(f \circ g)(t) = 1/2$ para todo $t \neq 0$ e $(f \circ g)(0) = 0$. Como g é evidentemente contínua, segue-se que f é descontínua na origem.

Uma noção de grande utilidade teórica é a de continuidade uniforme. Uma aplicação $f\colon X \to \mathbb{R}^n$, definida em $X \subset \mathbb{R}^m$, diz-se *uniformemente contínua* quando, para todo $\varepsilon > 0$ dado, pode-se obter $\delta > 0$ tal que $x, y \in X$, $|x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon$.

Por exemplo, toda aplicação Lipschitziana $f\colon X \to \mathbb{R}^n$ é uniformemente contínua: se $|f(x) - f(y)| \leq k|x - y|$ para $x, y \in X$ arbitrários dado $\varepsilon > 0$, tomamos $\delta = \varepsilon/k$. Então $|x - y| \leq \delta \Rightarrow |f(x) - f(y)| \leq k(\varepsilon/k) = \varepsilon$.

Assim, em particular, toda aplicação linear $A\colon \mathbb{R}^m \to \mathbb{R}^n$ é uniformemente contínua, o mesmo ocorrendo com a restrição de uma aplicação bilinear $\varphi\colon \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^p$ a um subconjunto limitado $X \subset \mathbb{R}^m \times \mathbb{R}^n$. A função $f\colon [0, +\infty) \to \mathbb{R}$, definida por $f(x) = \sqrt{x}$, é o exemplo clássico de função uniformemente contínua que não é Lipschitziana. (Vide Vol. 1, pág. 244.)

A composta $g \circ f$ de duas aplicações uniformemente contínuas g e f é ainda uniformemente contínua.

Uma aplicação $f \colon X \to \mathbb{R}^n$ é uniformemente contínua se, e somente se, suas funções coordenadas $f_1, \ldots, f_n \colon X \to \mathbb{R}$ o são.

(7.4). *Uma aplicação* $f \colon X \to \mathbb{R}^n$ *é uniformente contínua se, e somente se, para todo par de seqüências* (x_k), (y_k) *em* X, *com* $\lim(x_k - y_k) = 0$, *tem-se* $\lim[f(x_k) - f(y_k)] = 0$.

Com efeito, seja f uniformemente contínua e sejam (x_k), (y_k) seqüências em X, com $\lim(x_k - y_k) = 0$. Dado arbitrariamente $\varepsilon > 0$, existe $\delta > 0$ tal que $|x - y| < \delta, \quad x, y \in X \Rightarrow |f(x) - f(y)| < \varepsilon$. Correspondentemente a δ, existe $k_o \in \mathbb{N}$ tal que $k > k_o \Rightarrow |x_k - y_k| < \delta$, logo $k > k_0 \Rightarrow |f(x_k) - f(y_k)| < \varepsilon$. Assim $\lim[f(x_k) - f(y_k)] = 0$, o que prova que a condição é necessária.

Reciprocamente, se f não for uniformemente contínua, existe $\varepsilon > 0$ ao qual não corresponde um $\delta > 0$ que cumpra a exigência da definição. Neste caso, tomando sucessivamente $\delta = 1, \frac{1}{2}, \ldots, \frac{1}{k}, \ldots$ obtemos, para cada $k \in \mathbb{N}$, um par de pontos $x_k, y_k \in X$ com $|x_k - y_k| < 1/k$ e $|f(x_k) - f(y_k)| \geq \varepsilon$. Então $\lim(x_k - y_k) = 0$ mas não se tem $\lim[f(x_k) - f(y_k)] = 0$. Isto mostra a suficiência da condição.

Por exemplo, a função $f \colon \mathbb{R} \to \mathbb{R}$, dada por $f(x) = \cos(x^2)$, não é uniformente contínua. Basta tomar $x_k = \sqrt{(k+1)\pi}$ e $y_k = \sqrt{k\pi}$. Então

$$
\begin{aligned}
x_k - y_k &= \frac{(\sqrt{(k+1)\pi} - \sqrt{k\pi})(\sqrt{(k+1)\pi} + \sqrt{k\pi})}{\sqrt{(k+1)\pi} + \sqrt{k\pi}} = \\
&= \frac{\pi}{\sqrt{(k+1)\pi} + \sqrt{k\pi}} \,.
\end{aligned}
$$

Logo, $\lim(x_k - y_k) = 0$. Mas $\cos(x_k^2) = \cos(k+1)\pi = \pm 1$ e $\cos(y_k^2) = \cos(k\pi) = \mp 1$, portanto $|f(x_k) - f(y_k)| = 2$ para todo k. Assim não se tem $\lim[f(x_k) - f(y_k)] = 0$.

8 Homeomorfismos

Dados os conjuntos $X \subset \mathbb{R}^m$ e $Y \subset \mathbb{R}^n$, um *homeomorfismo* entre X e Y é uma bijeção contínua $f \colon X \to Y$, cuja inversa $f^{-1} \colon Y \to X$ também é contínua. Diz-se então que X e Y são conjuntos *homeomorfos*.

Uma bijeção $f \colon X \to Y$ pode muito bem ser contínua sem que sua inversa o seja. Um exemplo disso é a aplicação $f \colon [0, 2\pi) \to S^1$, do

intervalo semi-aberto $[0, 2\pi)$ sobre o círculo unitário $S^1 = \{(x, y) \in \mathbb{R}^2; x^2 + y^2 = 1\}$, definida por $f(t) = (\cos t, \operatorname{sen} t)$. Pelo Teorema 12, f é contínua. Além disso, f é evidentemente bijetiva. Mas sua inversa $f^{-1}\colon S^1 \to [0, 2\pi)$ é descontínua no ponto $p = (1, 0)$. Com efeito, para cada $k \in \mathbb{N}$ sejam $t_k = 2\pi - (1/k)$ e $z_k = f(t_k)$. Então $\lim_{k\to\infty} z_k = p$ mas não é verdade que $\lim_{k\to\infty} f^{-1}(z_k) = \lim_{k\to\infty} t_k$ seja igual a $f^{-1}(p) = 0$.

Um exemplo simples de homeomorfismo de $\mathbb{R}^n$ sobre si mesmo é dado por uma aplicação linear invertível $A\colon \mathbb{R}^n \to \mathbb{R}^n$. Com efeito, sua inversa $A^{-1}\colon \mathbb{R}^n \to \mathbb{R}^n$ é linear e portanto contínua.

A aplicação composta de dois homeomorfismos é um homeomorfismo. Também o inverso de um homeomorfismo é um homeomorfismo.

Os homeomorfismos desempenham na Topologia um papel análogo ao das congruências ou movimentos rígidos na Geometria Euclidiana: dois conjuntos homeomorfos são indistinguíveis do ponto de vista topológico.

Em seguida daremos alguns exemplos de homeomorfismos e de conjuntos homeomorfos.

A) *As translações* $T_a\colon \mathbb{R}^m \to \mathbb{R}^m$, $T_a(x) = x + a$. Com efeito, cada T_a é uma isometria, logo uma aplicação contínua. Como $(T_1)^{-1} = T_{-a}$, vemos que sua inversa também é continua.

B) *As homotetias* $H_\lambda\colon \mathbb{R}^m \to \mathbb{R}^m$, $H_\lambda(x) = \lambda \cdot x$, com $0 \neq \lambda \in \mathbb{R}$. Cada homotetia H_λ é uma transformação linear invertível, com $(H_\lambda)^{-1} = H_{\lambda^{-1}}$.

C) Duas bolas abertas quaisquer em $\mathbb{R}^m$ são homeomorfas, o mesmo ocorrendo com duas bolas fechadas arbitrárias em $\mathbb{R}^m$ ou duas esferas no mesmo espaço euclidiano. Com efeito, dadas $B(a; r)$ e $B(b; s)$, duas bolas abertas em $\mathbb{R}^m$, consideremos o homeomorfismo $\varphi\colon \mathbb{R}^m \to \mathbb{R}^m$, dado por $\varphi = T_b \circ H_{s/r} \circ T_{-a}$. Para cada $x \in \mathbb{R}^m$ temos $\varphi(x) = b + \dfrac{s}{r}(x - a)$. Isto mostra que φ consiste em primeiro transladar $B(a; r)$ de modo a pôr seu centro na origem, em seguida multiplicar todos os vetores (com origem 0) por s/r de modo que os vetores de comprimento $< r$ passem a ter comprimento $< s$ e, finalmente, transladar $B(0; s)$ de maneira a pôr seu centro no ponto b. Logo o homeomorfismo φ transforma $B(a; r)$ em $B(b; s)$ e, por restrição, define um homeomorfismo entre estas bolas. Note que $\varphi \mid B[a; r]$ também é um homeomorfismo sobre a bola fechada $B[b; s]$. Analogamente, φ transforma homeomorficamente a esfera $S[a; r]$ na esfera $S[b; s]$.

D) Toda bola aberta em $\mathbb{R}^m$ é homeomorfa ao espaço euclidiano $\mathbb{R}^m$. Como duas bolas abertas em $\mathbb{R}^m$ são sempre homeomorfas, basta estabelecer um homeomorfismo entre $\mathbb{R}^m$ e a bola unitária $B = B(0;1)$, de centro na origem e raio 1. Sejam $f\colon \mathbb{R}^m \to B$ e $g\colon B \to \mathbb{R}^m$ definidas por $f(x) = x/(1+|x|)$ e $g(y) = y/(1-|y|)$. Evidentemente, f e g são contínuas. Verifica-se facilmente que $g(f(x)) = x$ e $f(g(y)) = y$ para quaisquer $x \in \mathbb{R}^m$ e $y \in B$. Logo f e g são homeomorfismos, um sendo o inverso do outro.

E) Seja $f\colon X \to \mathbb{R}^n$ uma aplicação contínua, definida no conjunto $X \subset \mathbb{R}^m$. Seu *gráfico* é o conjunto $G \subset \mathbb{R}^m \times \mathbb{R}^n = \mathbb{R}^{m+n}$, formado pelos pontos $(x, f(x))$, onde $x \in X$. Afirmamos que o domínio X e o gráfico G da aplicação contínua f são homeomorfos. Basta considerar $\tilde{f}\colon X \to G$, definida por $\tilde{f}(x) = (x, f(x))$. Pelo Corolário do Teorema 12, $\tilde{f}$ é contínua. Sua inversa $g\colon G \to X$, dada por $g(x, f(x)) = x$, é contínua pois é a restrição a G da projeção $\pi_1\colon \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^m$. Como caso particular deste exemplo temos: $\mathbb{R} - \{0\}$ é homeomorfo à hipérbole $H = \{(x,y) \in \mathbb{R}^2; xy = 1\}$ pois H é o gráfico da função contínua $f\colon \mathbb{R} - \{0\} \to \mathbb{R}$, dada por $f(x) = 1/x$. Outro caso particular: seja $S^m = \{y \in \mathbb{R}^{m+1}; \langle y, y\rangle = 1\}$ a esfera unitária m-dimensional. Seu *hemisfério norte* é o conjunto $S^m_+ = \{y \in S^m; y_{m+1} > 0\}$. Ele é homeomorfo à bola aberta unitária $B = \{x \in \mathbb{R}^m; |x| < 1\}$ do espaço $\mathbb{R}^m$, pois S^m_+ é o gráfico da função contínua $f\colon B \to \mathbb{R}$, onde $f(x) = (1-|x|^2)^{1/2}$. Com efeito, tem-se $y \in S^m_+ \Leftrightarrow y = (x,t)$, com $x \in B$ e $t^2 = 1 - |x|^2, \quad t > 0$, ou seja, $t = (1-|x|^2)^{1/2}$.

F) *A projeção estereográfica.* Sejam $S^m = \{x \in \mathbb{R}^{m+1}; \langle x, x\rangle = 1\}$ a esfera unitária m-dimensional e $p = (0, \ldots, 0, 1) \in S^m$ seu pólo norte. A projeção estereográfica $\varphi\colon S^m - \{p\} \to \mathbb{R}^m$ é um homeomorfismo entre a esfera menos o pólo norte e o espaço euclidiano $\mathbb{R}^m$. Geometricamente, $\varphi(x)$ é o ponto em que a semi-reta $px \subset \mathbb{R}^{m+1}$ corta o hiperplano $x_{m+1} = 0$, o qual identificamos com $\mathbb{R}^m$. A fim de obter uma fórmula para φ, notemos que os pontos da semi-reta px são da forma $p + t(x-p)$, com $t > 0$. Este ponto está no hiperplano $x_{m+1} = 0$ quando sua última coordenada, $1 + t(x_{m+1} - 1)$ é zero. Daí tiramos $t = 1/(1-x_{m+1})$ e por conseguinte $\varphi(x) = x'/(1-x_{m+1})$, onde convencionamos escrever $x' = (x_1, \ldots, x_m)$ quando $x = (x_1, \ldots, x_m, x_{m+1})$. Vemos então que $\varphi\colon S^m - \{p\} \to \mathbb{R}^m$ é contínua. Se definirmos $\xi\colon \mathbb{R}^m \to S^m - \{p\}$ pondo $\xi(y) = x$ onde, na notação acima, $x' = 2y/(|y|^2+1)$ e $x_{m+1} = (|y|^2-1)/(|y|^2+1)$, constataremos que ξ é contínua, que $\varphi(\xi(y)) = y$ para todo $y \in \mathbb{R}^m$ e

$\xi(\varphi(x)) = x$ para todo $x \in S^m - \{p\}$. Logo φ é um homeomorfismo e ξ o seu homeomorfismo inverso. A projeção estereográfica possui várias propriedades geométricas interessantes, especialmente no caso $m = 2$. Este foi o caso descoberto por Gauss, para mostrar que a esfera unitária S^2 pode ser interpretada como o plano $\mathbb{R}^2$ ao qual se acrescentou um "ponto no infinito".

G) Dado arbitrariamente $a \in \mathbb{R}$, seja $p = (\cos a, \operatorname{sen} a)$. A aplicação $\xi\colon (a, a+2\pi) \to S^1 - \{p\}$, definida por $\xi(t) = (\cos t, \operatorname{sen} t)$, é um homeomorfismo. Com efeito, ξ é contínua, bijetiva e, se uma seqüência de pontos $t_n \in (a, a+2\pi)$ não possui valores de aderência nesse intervalo então seus valores de aderência em $\mathbb{R}$ só podem ser a e/ou $a + 2\pi$. Nestas condições, tem-se $\lim \xi(t_n) = p$. Segue-se daí que, dada uma seqüência de pontos $s_n \in (a, a+2\pi)$ com $\lim \xi(s_n) = q \in S^1 - \{p\}$, qualquer valor de aderência de (s_n) é um número $b \in (a, a+2\pi)$ o qual cumpre (pela continuidade de ξ) a condição $\xi(b) = q$. Como ξ é injetiva, b é o único valor de aderência de (s_n), logo $\lim s_n = b$. Em suma, $\lim \xi(s_n) = q \in S^1 - \{p\} \Rightarrow \lim s_n = b = \xi^{-1}(q)$. Isto diz que $\xi^{-1}\colon S^1 - \{p\} \to (a, a+2\pi)$ é contínua, ou seja, que ξ é um homeomorfismo.

9 Limites

Sejam $f\colon X \to \mathbb{R}^n$ uma aplicação definida num conjunto $X \subset \mathbb{R}^m$ e $a \in \mathbb{R}^m$ um ponto de acumulação de X.

Diz-se que o ponto $b \in \mathbb{R}^n$ é o *limite de $f(x)$ quando x tende para a*, e escreve-se

$$b = \lim_{x \to a} f(x),$$

para significar o seguinte:

Dado qualquer $\varepsilon > 0$, pode-se obter $\delta > 0$ tal que $x \in X$, $0 < |x - a| < \delta \Rightarrow |f(x) - b| < \varepsilon$.

Para que tenha sentido a afirmação $\lim_{x \to a} f(x) = b$, não é necessário que a pertença a X, ou seja, que f seja definida no ponto a. Pelo contrário, mesmo que a pertença ao domínio de f, o valor $f(a)$ não desempenha papel algum na definição de limite: importam apenas os valores $f(x)$ para x próximo, porém diferente, de a.

Quando existe, o limite é *único*: se $a \in X'$ e se tem, ao mesmo tempo, $\lim_{x \to a} f(x) = b$, $\lim_{x \to a} f(x) = c$, então $b = c$. Com efeito, dado

arbitrariamente $\varepsilon > 0$, existe $\delta > 0$ tal que $x \in X$, $0 < |x-a| < \delta \Rightarrow |f(x)-b| < \varepsilon/2$, $|f(x)-c| < \varepsilon/2$. Como $a \in X'$, podemos achar $x \in X$ tal que $0 < |x-a| < \delta$. Usando este x, temos: $|b-c| < |b-f(x)| + |f(x)-c| < \varepsilon$. Sendo ε arbitrário, resulta $b = c$.

A continuidade se exprime em termos de limite: se o ponto $a \in X$ é isolado então toda aplicação $f\colon X \to \mathbb{R}^n$ é contínua no ponto a. Se, porém, $a \in X'$ então f é contínua no ponto a se, e somente se, $\lim\limits_{x\to a} f(x) = f(a)$.

(9.1) *Para que seja* $\lim\limits_{x\to a} f(x) = b$, *é necessário e suficiente que* $\lim\limits_{x\to a} f(x_k) = b$, *seja qual for a seqüência de pontos* $x_k \in X - \{a\}$ *com* $\lim x_k = a$.

A afirmação acima se demonstra da maneira análoga ao Teorema 13. Nas aplicações, a seguinte conseqüência de (9.1) é freqüentemente útil: (Compare com o Vol. 1, pág. 200, Corolário 2.)

Para que exista $\lim\limits_{x\to a} f(x)$ *é suficiente que exista* $\lim f(x_k)$ *seja qual for a seqüência de pontos* $x_k \in X - \{a\}$ *com* $\lim x_k = a$.

Quando $X \subset \mathbb{R}$ e a é ponto de acumulação à esquerda, é bastante tomar a seqüência (x_k) crescente. (Ou decrescente, se a é ponto de acumulação à direita.)

Vale para limites o análogo do Teorema 12:

(9.2) *Seja a um ponto de acumulação do conjunto $X \subset \mathbb{R}^m$. Dada a aplicação $f\colon X \to \mathbb{R}^n$, cujas funções coordenadas são $f_1, \ldots, f_n\colon X \to \mathbb{R}$, tem-se $\lim\limits_{x\to a} f(x) = b = (b_1, \ldots, b_n)$ se, e somente se, $\lim\limits_{x\to a} f_i(x) = b_i$ para cada $i = 1, 2, \ldots, n$.*

A demonstração é praticamente a mesma do Teorema 12.

A noção de limite se relaciona com as operações do espaço vetorial $\mathbb{R}^m$ do seguinte modo:

(9.3) Sejam $X \subset \mathbb{R}^m$, $a \in X'$, $b, c \in \mathbb{R}^n$, $f, g\colon X \to \mathbb{R}^n$ e $\alpha\colon X \to \mathbb{R}$ tais que $\lim\limits_{x\to a} f(x) = b$, $\lim\limits_{x\to a} g(x) = c$, $\lim\limits_{x\to a} \alpha(x) = \alpha_0$. Então:

1) $\lim\limits_{x\to a} (f(x) + g(x)) = b + c$;

2) $\lim\limits_{x\to a} \alpha(x) \cdot f(x) = \alpha_0 \cdot b$;

3) $\lim\limits_{x\to a} \langle f(x), g(x)\rangle = \langle b, c\rangle$;

4) Além disso, seja $\varphi\colon \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^p$ bilinear. Dadas agora $f\colon X \to \mathbb{R}^m$, $g\colon X \to \mathbb{R}^n$, com $\lim\limits_{x\to a} f(x) = 0$ e g limitada então $\lim\limits_{x\to a} \varphi(f(x), g(x)) = 0$.

Como casos particulares, temos $\lim_{x\to a} \langle f(x), g(x)\rangle = 0$ e $\lim_{x\to a} \alpha(x)\cdot f(x) = 0$ se um dos fatores é limitado e o outro tende para zero.

As três primeiras afirmações decorrem imediatamente da caracterização do limite por meio de seqüências, juntamente com o corolário do Teorema 4. A afirmação 4) resulta de ser $|\varphi(f(x), g(x))| \leq M|f(x)|\,|g(x)|$ para uma constante $M > 0$ que depende apenas da aplicação bilinear φ.

Como exemplo, consideremos $f\colon \mathbb{R}^2 - \{0\} \to \mathbb{R}$, definida por $f(x,y) = x^2y/(x^2+y^2)$. Mostremos que $\lim_{(x,y)\to(0,0)} f(x,y) = 0$. Como $f(x,y)$ é o produto de x por $xy/(x^2+y^2)$ e, evidentemente, $\lim_{(x,y)\to(0,0)} x = 0$, resta mostrar que $xy/(x^2+y^2)$ é limitada. Mas, para $(x,y) \neq (0,0)$, $xy/(x^2+y^2) = \big(x/\sqrt{x^2+y^2}\big)\big(y/\sqrt{x^2+y^2}\big) = \cos\theta \cdot \operatorname{sen}\theta$, onde θ é o ângulo que o vetor $z = (x,y)$ forma com o eixo das abcissas. Logo $xy/(x^2+y^2)$ tem valor absoluto ≤ 1.

A relação entre o limite e a composição de aplicações é a seguinte: sejam $f\colon X \to \mathbb{R}^n$, $g\colon Y \to \mathbb{R}^p$, $a \in X'$, $b \in Y'$ e $f(X) \subset Y$.

1) Se $\lim_{x\to a} f(x) = b$ e g é contínua no ponto b então $\lim_{x\to a} g(f(x)) = g(b)$.

2) Se $\lim_{x\to a} f(x) = b$, $\lim_{y\to b} g(y) = c$ e $x \neq a \Rightarrow f(x) \neq b$ então $\lim_{x\to a} g(f(x)) = c$.

A afirmação 1) fornece mais uma demonstração para as propriedades operatórias dos limites vistas há pouco. Outra conseqüência é a seguinte: $\lim_{x\to a} f(x) = b$ então $\lim_{x\to a} |f(x)| = |b|$.

Usando a afirmação 2), concluímos que se existe $\lim_{x\to a} f(x) = b$ então, para qualquer vetor $u \neq 0$ tem-se $\lim_{t\to 0} f(a+tu) = b$. Segue-se daí que não existe $\lim_{(x,y)\to(0,0)} \dfrac{xy}{x^2+y^2}$ pois, se pusermos $u = (\alpha,\beta)$, vem $\lim_{t\to 0} f(t\alpha, t\beta) = \dfrac{\alpha\beta}{\alpha^2+\beta^2}$ e este limite varia com α e β.

Além das regras operacionais com limites, a manipulação formal do símbolo "lim"é muito facilitada pelo princípio de *permanência das desigualdades*, que mostraremos agora. Trata-se do seguinte:

(9.4) *Sejam $f, g\colon X \to \mathbb{R}$ definidas em $X \subset \mathbb{R}^m$ e seja $a \in X'$. Suponhamos que $f(x) \leq g(x)$ para todo $x \in X - \{a\}$. Se existirem $\lim_{x\to a} f(x) = b$ e $\lim_{x\to a} g(x) = c$, deverá ser $b \leq c$.*

Com efeito, do contrário teríamos $c < b$ e portanto $\varepsilon = (b-c)/2$ seria positivo. Correspondendo a este ε, poderíamos obter $\delta > 0$ tal que $x \in X, \quad 0 < |x-a| < \delta \Rightarrow f(x) \in (b-\varepsilon, b+\varepsilon)$ e $g(x) \in (c-\varepsilon, c+\varepsilon)$. Como $b - \varepsilon = c + \varepsilon$, isto obrigaria $g(x) < f(x)$ para todo esses valores de x, uma contradição.

Os fatos mais importantes que desejamos discutir sobre limites serão estabelecidos agora.

Teorema 14. *Seja $f\colon X \to \mathbb{R}^n$ uniformemente contínua no conjunto $X \subset \mathbb{R}^m$. Então, para todo $a \in X'$, existe $\lim_{x\to a} f(x)$.*

Demonstração: Sendo f uniformemente contínua, toda seqüência de Cauchy de pontos $x_k \in X$ é transformada por f numa seqüência de Cauchy de pontos $f(x_k) \in \mathbb{R}^n$. Em particular, para toda seqüência de pontos $x_k \in X - \{a\}$ com $\lim_{k\to\infty} x_k = a$, existe $\lim_{k\to\infty} f(x_k) = b$. Este valor b não depende da seqüência escolhida, pois se tivéssemos outra escolha de pontos $y_k \in X - \{a\}$ tal que $\lim_{k\to\infty} f(y_k) = c$, com $c \neq b$, então a seqüência de pontos $z_k \in X - \{a\}$ tal que $z_{2k} = x_k$ e $z_{2k-1} = y_k$ ainda cumpriria $\lim z_k = a$, mas agora $\lim_{k\to\infty} f(z_k)$ não existiria em virtude de $(f(z_k))$ possuir duas subseqüências convergindo para limites distintos b e c. Segue-se da caracterização do limite por meio de seqüências que se tem $\lim_{x\to a} f(x) = b$.

Assim, por exemplo, a função contínua $f\colon \mathbb{R}^2 - \{0\} \to \mathbb{R}$, definida por $f(x,y) = xy/(x^2+y^2)$, não é uniformemente contínua em qualquer conjunto $X \subset \mathbb{R}^2$ do qual a origem seja um ponto de acumulação, pois não existe $\lim_{(x,y)\to(0,0)} f(x,y)$.

Corolário. *Dada $f\colon X \to \mathbb{R}^n$ uniformemente contínua, seja $\overline{X} = X \cup X'$. Existe uma única aplicação uniformemente contínua $\bar{f}\colon \overline{X} \to \mathbb{R}^n$ tal que $\bar{f} \mid X = f$.*

Para todo ponto $\bar{x} \in X' - X$, ponhamos $\bar{f}(\bar{x}) = \lim_{x\to\bar{x}} f(x)$. (Este limite existe, pelo teorema que acabamos de demonstrar.) Se $x \in X$, pomos $\bar{f}(x) = f(x)$. Fica assim definida uma aplicação $\bar{f}\colon \overline{X} \to \mathbb{R}^n$ que estende f. Afirmamos que $\bar{f}$ é uniformemente contínua. Com efeito, dado $\varepsilon > 0$ existe, pela continuidade uniforme de f, um número $\delta > 0$ tal que $x, y \in X, \quad |x-y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon/2$. Este é o δ que corresponde a ε para provar a continuidade uniforme de $\bar{f}$. De fato, sejam $\bar{x}, \bar{y} \in \overline{X}$ tais que $|\bar{x} - \bar{y}| < \delta$. Temos $\bar{x} = \lim x_k$ e $\bar{y} = \lim y_k$

com $x_k, y_k \in X$ para todo $k \in \mathbb{N}$. (Se já for $\bar{x} \in X$, basta tomar $x_k = x$ para todo k. Se, porém, $\bar{x} \in X' - X$, teremos $x_k \neq \bar{x}$ para todo k. O mesmo quanto a $\bar{y}$.) Como $|\bar{x} - \bar{y}| < \delta$, segue-se que $|x_k - y_k| < \delta$ para todo k suficientemente grande, digamos $k > k_0$. Então $k > k_0 \Rightarrow |f(x_k) - f(y_k)| < \varepsilon/2$. Daí $|\bar{f}(\bar{x}) - \bar{f}(\bar{y})| = |\lim f(x_k) - \lim f(y_k)| = \lim |f(x_k) - f(y_k)| \leq \varepsilon/2 < \varepsilon$.

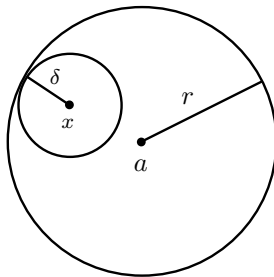
10 Conjuntos abertos

Seja X um subconjunto do espaço euclidiano $\mathbb{R}^n$. Um ponto $a \in X$ chama-se um *ponto interior* a X quando é centro de alguma bola aberta contida em X, ou seja, quando existe $\delta > 0$ tal que $|x-a| < \delta \Rightarrow x \in X$. O *interior* de X é o conjunto $\operatorname{int} X$, formado pelos pontos interiores a X. Quando $x \in \operatorname{int} V$, dizemos que o conjunto V é uma *vizinhança* do ponto x.

Dizer que um ponto $a \in X$ não é interior a X equivale a afirmar que toda bola aberta de centro a contém pontos do complementar de X, ou seja, que, para todo $\delta > 0$ existe $y \in \mathbb{R}^n - X$ com $|y - a| < \delta$.

Um conjunto $X \subset \mathbb{R}^n$ chama-se *aberto* quando todos os seus pontos são interiores, isto é, quando para cada $x \in X$ existe $\delta > 0$ tal que $B(x;\delta) \subset X$. Assim, X é aberto $\Leftrightarrow \operatorname{int} X = X$.

(10.1) *Uma bola aberta $B(a;r) \subset \mathbb{R}^n$ é um exemplo de conjunto aberto.*

Com efeito, dado qualquer $x \in B(a;r)$, temos $|x - a| < r$, logo o número $\delta = r - |x - a|$ é positivo. Afirmamos que $B(x;\delta) \subset B(a;r)$. De fato $y \in B(x;\delta) \Rightarrow |y-a| \leq |y-x|+|x-a| < \delta+|x-a| = r \Rightarrow y \in B(a;r)$.

Também é aberto em $\mathbb{R}^n$ o conjunto $X = \mathbb{R}^n - B[a;r]$, complementar da bola fechada $B[a;r]$. Temos $X = \{x \in \mathbb{R}^n; |x - a| > r\}$. Dado arbitrariamente $x \in X$, seja $\delta = |x-a|-r$. Afirmamos que $B(x;\delta) \subset X$.

Com efeito, $y \in B(x;\delta) \Rightarrow |x-a| \leq |x-y| + |y-a| < \delta + |y-a| = |x-a| - r + |y-a| \Rightarrow |y-a| > r \Rightarrow y \in X$.

(10.2) *Para todo conjunto* $X \subset \mathbb{R}^n$, $\operatorname{int} X$ *é um conjunto aberto.*

De fato, se $a \in \operatorname{int} X$ então existe $r > 0$ tal que $B(a;r) \subset X$. Se $x \in B(a;r)$ então pondo $\delta = r - |x-a|$, vemos que $B(x;\delta) \subset B(a;r)$, donde $B(x;\delta) \subset X$ e portanto $x \in \operatorname{int} X$. Assim, todo ponto $a \in \operatorname{int} X$ é centro de bola $B(a;r)$ contida em $\operatorname{int} X$, o que prova que $\operatorname{int} X$ é aberto.

Por outro lado, uma bola fechada $B[a;r]$ em $\mathbb{R}^n$ não é um conjunto aberto pois, se tomarmos arbitrariamente um vetor unitário $u \in \mathbb{R}^n$, o ponto $x = a + r \cdot u$ é tal que $|x-a| = r$, logo $x \in B[a;r]$. Mas nenhuma bola aberta $B(x;\delta)$ está contida em $B[x;r]$. Com efeito, tomando $y = a + (r+\delta/2)u$, temos $|y-x| = \delta/2 < \delta$ mas $|y-a| = r + \delta/2 > r$. Assim, $y \in B(x,\delta)$ mas $y \notin B[a;r]$. Este argumento mostra, de fato, que os pontos da esfera $S[a;r]$ não são interiores à bola fechada de centro a e raio r. Portanto, $\operatorname{int} B[a;r] = B(a;r)$.

Dados um conjunto X e um ponto $a \in \mathbb{R}^n$, há três possibilidades que se excluem mutuamente: ou $a \in \operatorname{int} X$, ou $a \in \operatorname{int}(\mathbb{R}^n - X)$ ou então toda bola aberta de centro a contém pontos de X e pontos do complementar de X. Os pontos com esta última propriedade constituem ∂X, que chamaremos a *fronteira* de X. Os pontos $y \in \partial X$ são chamados *pontos-fronteira* de X.

Assim, por exemplo, se X é a bola fechada $B[a;r]$, temos $\partial X = S[a;r]$ = esfera de centro a e raio r. Observe que se chamarmos de Y a bola aberta $B(a;r)$ teremos $\partial X = \partial Y$.

Um conjunto $A \subset \mathbb{R}^n$ é aberto se, e somente se, nenhum dos seus pontos é ponto fronteira de A, ou seja, se, e somente se, $A \cap \partial A = \varnothing$.

Teorema 15. *Os conjuntos abertos do espaço euclidiano* $\mathbb{R}^n$ *gozam das seguintes propriedades:*

1) *O conjunto vazio* $\varnothing$ *e o espaço* $\mathbb{R}^n$ *inteiro são abertos;*

2) *A interseção* $A = A_1 \cap \cdots \cap A_k$ *de um número finito de conjuntos abertos* $A_1, \ldots, A_k$ *é um conjunto aberto;*

3) *A reunião* $A = \bigcup_{\lambda \in L} A_\lambda$ *de uma família qualquer* $(A_\lambda)_{\lambda \in L}$ *de conjuntos abertos* A_λ *é um conjunto aberto.*

Demonstração: Um conjunto só pode deixar de ser aberto se contiver algum ponto que não seja interior. Como $\varnothing$ não contém ponto algum,

é aberto. $\mathbb{R}^n$ é obviamente aberto. Para provar 2), seja $a \in A$. Então, para cada $i = 1, \ldots, k$, temos $a \in A_i$. Como A_i é aberto, existe $\delta_i > 0$ tal que $B(a;\delta_i) \subset A_i$. Seja $\delta = \min\{\delta_1, \ldots, \delta_k\}$. Então $B(a;\delta) \subset A_i$ para cada i, donde $B(a;\delta) \subset A$. Finalmente, provemos 3). Dado $a \in A$, existe $\lambda \in L$ tal que $a \in A_\lambda$. Sendo A_λ aberto, existe $\delta > 0$ com $B(a;\delta) \subset A_\lambda \subset A$. Logo A é aberto.

Fixemos um conjunto $X \subset \mathbb{R}^n$. Um subconjunto $A \subset X$ diz-se *aberto em* X quando, para cada $a \in A$ existe $\delta > 0$ tal que $B(a;\delta) \cap X \subset A$. Noutras palavras, para cada $a \in A$ existe $\delta > 0$ tal que os pontos x, *pertencentes a* X, que cumprem a condição $|x - a| < \delta$ estão em A.

Por exemplo, $A = (0, 1]$ é aberto em $X = [0, 1]$.

Chega-se à noção de conjunto aberto em X quando se faz abstração dos demais pontos do espaço $\mathbb{R}^n$, considerando-se apenas os pontos de X, e se procura então imitar a definição de conjunto aberto.

Como se vê facilmente, quando $X \subset \mathbb{R}^n$ é aberto, um subconjunto $A \subset X$ é aberto em X se, e somente se, é aberto no sentido usual de $\mathbb{R}^n$.

(10.3) *Mais geralmente, um conjunto $A \subset X$ é aberto em X se, e somente se, existe um aberto $B \subset \mathbb{R}^n$ tal que $A = X \cap B$.*

Com efeito, se for A aberto em X, tome B igual à reunião das bolas $B(a;\delta)$ com centro nos pontos $a \in A$, tais que $B(a;\delta) \cap X \subset A$. Reciprocamente, se for $A = X \cap B$, com B aberto em $\mathbb{R}^n$, para cada $a \in A$ existe uma bola $B(a;\delta) \subset B$ logo $B(a;\delta) \cap X \subset B \cap X = A$ portanto A é aberto em X.

Vale para os abertos em X um resultado análogo ao do Teorema 15: $\varnothing$ e X são abertos em X; uma interseção finita e uma reunião qualquer de abertos em X é ainda um conjunto aberto em X.

O interesse desta nova noção reside no teorema abaixo.

Teorema 16. *Seja $f\colon X \to \mathbb{R}^n$ uma aplicação definida no conjunto $X \subset \mathbb{R}^n$. A fim de que f seja contínua, é necessário e suficiente que a imagem inversa $f^{-1}(A)$ de todo aberto $A \subset \mathbb{R}^n$ seja um conjunto aberto em X.*

Demonstração: (Necessário.) Se f é contínua e $A \subset \mathbb{R}^n$ é aberto, tomemos um ponto $a \in f^{-1}(A)$. Então $f(a) \in A$. Pela definição de aberto, existe $\varepsilon > 0$ tal que $B(f(a);\varepsilon) \subset A$. Sendo f contínua, existe $\delta > 0$ tal que $x \in X$, $|x - a| < \delta \Rightarrow |f(x) - f(a)| < \varepsilon$. Isto significa que $f(B(a;\delta) \cap X) \subset B(f(a);\varepsilon) \subset A$, donde $B(a;\delta) \cap X \subset f^{-1}(A)$. Logo $f^{-1}(A)$ é aberto em X.

(Suficiente.) Se a imagem inversa por f de todo aberto de $\mathbb{R}^n$ é aberto em X; então, dados $a \in X$ e $\varepsilon > 0$, como $B(f(a);\varepsilon)$ é aberto, concluímos que $A = \{x \in X; |f(x) - f(a)| < \varepsilon\}$ é aberto em X. Evidentemente, $a \in A$. Logo existe $\delta > 0$ tal que $B(a;\delta) \cap X \subset A$. Isto significa porém que $x \in X$, $|x - a| < \delta \Rightarrow |f(x) - f(a)| < \varepsilon$, ou seja, que f é contínua no ponto a. Como $a \in X$ é qualquer, f é contínua.

Observação: O mesmo resultado continuaria válido se substituíssemos, no enunciado acima, a expressão "todo aberto $A \subset \mathbb{R}^n$"por "todo conjunto $A \subset f(X)$, aberto em $f(X)$". Com efeito, um aberto em $f(X)$ tem a forma $A' \cap f(X)$ onde A' é aberto em $\mathbb{R}^n$ e $f^{-1}(A' \cap f(X)) = f^{-1}(A')$. Logo, este novo enunciado, embora aparentemente mais geral, reduz-se ao que foi demonstrado no Teorema 16.

Como aplicação do teorema acima, vemos que se $f\colon \mathbb{R}^n \to \mathbb{R}$ é uma função contínua então, para todo número real a, o conjunto $A = \{x \in \mathbb{R}^n; f(x) < a\}$ é aberto, pois A é a imagem inversa do intervalo aberto $(-\infty, a)$ pela função f. Usando também o fato de que a interseção finita de abertos é ainda um aberto, concluimos que, dadas k funções contínuas $f_1, \ldots, f_k\colon \mathbb{R}^n \to \mathbb{R}$, o conjunto $A = \{x \in \mathbb{R}^n; f_1(x) < a_1, \ldots, f_k(x) < a_k\}$ é aberto. A mesma conclusão vale se as funções $f_1, \ldots, f_k$ forem definidas num subconjunto aberto $X \subset \mathbb{R}^n$.

Reobtemos assim o fato de que a bola $B(a;r)$ é um subconjunto aberto de $\mathbb{R}^n$. Com efeito, a função $x \mapsto |x - a|$ é contínua e $B(a;r)$ é o conjunto dos pontos de $\mathbb{R}^n$ onde esta função assume valores $< r$.

(10.4) *Se $A \subset \mathbb{R}^m$ e $B \subset \mathbb{R}^n$ são abertos então o produto cartesiano $A \times B \subset \mathbb{R}^m \times \mathbb{R}^n = \mathbb{R}^{m+n}$ é aberto.*

Com efeito, considerando as projeções $\pi_1\colon \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^m$ e $\pi_2\colon \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^n$, que são aplicações contínuas, temos $A \times B = \pi_1^{-1}(A) \cap \pi_2^{-1}(B)$. Evidentemente se tivermos $A_1 \subset \mathbb{R}^{m_1}, \ldots, A_k \subset \mathbb{R}^{m_k}$ abertos, o produto cartesiano $A_1 \times \cdots \times A_k \subset \mathbb{R}^{m_1 + \cdots + m_k}$ será aberto por uma razão semelhante.

Dados $X \subset \mathbb{R}^m$, $Y \subset \mathbb{R}^n$, uma aplicação $f\colon X \to Y$ diz-se *aberta* quando, para cada $A \subset X$ aberto em X, sua imagem $f(A)$ é um subconjunto aberto em Y.

(10.5) *Cada uma das m projeções $\pi_i\colon \mathbb{R}^m \to \mathbb{R}$ é uma aplicação aberta.*

Isto fica mais fácil de ver se usarmos em $\mathbb{R}^m$ a norma do máximo, em relação à qual uma bola aberta de centro $a = (a_1, \ldots, a_m)$ é um produto

cartesiano $B(a;\delta) = \prod_{j=1}^{m}(a_j - \delta, a_j + \delta)$.

Seja então $A \subset \mathbb{R}^m$ aberto. Para provar que $\pi_i(A)$ é aberto em $\mathbb{R}$, tomamos $a_i = \pi_i(a) \in \pi_i(A)$, $a \in A$. Como A é aberto, existe $\delta > 0$ tal que $B = \prod_{j=1}^{m}(a_j - \delta, a_j + \delta) \subset A$. Então $(a_i - \delta, a_i + \delta) = \pi_i(B) \subset \pi_i(A)$, logo $\pi_i(A)$ é aberto.

11 Conjuntos fechados

Um ponto $a \in \mathbb{R}^n$ diz-se *aderente* a um conjunto $X \subset \mathbb{R}^n$ quando é limite de uma seqüência de pontos desse conjunto.

Por exemplo, todo ponto a pertencente a X é aderente a X pois podemos escrever $a = \lim x_k$, com $x_k = a$ para todo $k \in \mathbb{N}$. Mas a pode ser aderente a X sem pertencer a X; neste caso, a é necessariamente um ponto de acumulação do conjunto X. Por exemplo, se $X = B(0;1) \subset \mathbb{R}^n$ é a bola aberta de centro na origem e raio 1 em $\mathbb{R}^n$, o ponto $e_1 = (1,0,\ldots,0)$ não pertence a X. Mas, pondo $x_k = (1 - 1/k, 0, \ldots, 0)$, vemos que $x_k \in X$ para todo $k \in \mathbb{N}$ e $\lim x_k = e_1$, logo e_1 é aderente a X.

A fim de que o ponto a seja aderente ao conjunto X, é necessário e suficiente que toda bola aberta de centro a contenha algum ponto de X. Com efeito, a condição é necessária em virtude da definição de limite de uma seqüência e é suficiente porque, se ela se verifica, em cada bola $B(a;1/k)$ podemos escolher um ponto $x_k \in X$ e assim obtemos uma seqüência com $\lim x_k = a$.

O conjunto dos pontos aderentes a X chama-se o *fecho* de X e é indicado com a notação $\overline{X}$.

Pelo que vimos acima, a fim de que um ponto $b \in \mathbb{R}^n$ *não pertença* ao fecho de X, é necessário e suficiente que exista uma bola aberta de centro b que não contém pontos de X. Noutros termos: $b \in \complement\overline{X} \Leftrightarrow \exists r > 0;$ $B(b;r) \cap X = \varnothing$.

Como toda bola aberta é um conjunto aberto e todo aberto que contém um ponto contém também uma bola aberta com centro nesse ponto, as condições acima podem ser reformuladas com abertos, em vez de bolas:

1. Tem-se $a \in \overline{X}$ se, e somente se, todo aberto que contém a intersecta o conjunto X. (Isto é, A aberto, $a \in A \Rightarrow A \cap X \neq \varnothing$.)

2. Tem-se $b \notin \overline{X}$ se, e somente se, existe um aberto contendo b e disjunto de X. (Isto é, existe A aberto com $b \in A$ e $A \cap X = \varnothing$.)

O fecho de uma bola aberta $B(a;r)$ é a bola fechada $B[a;r]$. Se $X = \mathbb{Q}^n$ é o conjunto dos pontos de $\mathbb{R}^n$ cujas coordenadas são números racionais, então $\overline{X} = \mathbb{R}^n$.

Um conjunto $X \subset \mathbb{R}^n$ chama-se *fechado* quando contém todos seus pontos aderentes, isto é, quando $X = \overline{X}$.

Dizer que $X \subset \mathbb{R}^n$ é fechado significa, portanto, o seguinte: se $\lim x_k = a$ e $x_k \in X$ para todo $k \in \mathbb{N}$, então $a \in X$.

Por exemplo, uma bola fechada $B[a;r]$ é um subconjunto fechado do espaço $\mathbb{R}^n$ pois se $|x_k - a| \le r$ para todo k e $\lim x_k = b$ então $|b - a| = \lim |x_k - a| \le r$.

Daí resulta que o fecho de todo conjunto limitado $X \subset \mathbb{R}^n$ é limitado. Com efeito, temos $X \subset B$, onde B é uma bola fechada. Logo $\overline{X} \subset \overline{B} = B$, donde $\overline{X}$ é limitado.

Para todo $X \subset \mathbb{R}^n$, o complementar do fecho de X é um aberto. Com efeito, seja $A = \complement\, \overline{X}$. Para todo $b \in A$ existe $r > 0$ tal que $B(b;r) \cap X = \varnothing$. Afirmamos que $B(b;r) \subset A$. Com efeito, se $y \in B(b;r)$ então $B(b;r)$ é um aberto contendo y e disjunto de X, logo $y \in \complement\, \overline{X} = A$.

Em particular, se $X \subset \mathbb{R}^n$ é um conjunto fechado, seu complementar $\complement\, X$ é aberto em $\mathbb{R}^n$ (pois $\complement\, X = \complement\, \overline{X}$).

Reciprocamente, se $X \subset \mathbb{R}^n$ é tal que $A = \complement\, X$ é um conjunto aberto então $y \notin X \Rightarrow y \in A \Rightarrow B(y;r) \subset A$ para algum $r > 0 \Rightarrow B(y;r) \cap X = \varnothing \Rightarrow y$ não é aderente a X. Assim, todo ponto aderente a X deve pertencer a X e conseqüentemente X é fechado.

Acabamos portanto de demonstrar o

Teorema 17. *Um conjunto é fechado se, e somente se, seu complementar é aberto.*

Corolário. *O fecho de todo conjunto é um conjunto fechado.*

O Corolário acima diz que $\overline{\overline{X}} = \overline{X}$ para todo $X \subset \mathbb{R}^n$.

O Teorema 17 faz com que o resultado seguinte seja uma conseqüência imediata do Teorema 15.

Teorema 18. *Os conjuntos fechados do espaço euclidiano $\mathbb{R}^n$ gozam das seguintes propriedades:*

1) *O conjunto vazio $\varnothing$ e o espaço inteiro $\mathbb{R}^n$ são fechados;*

2) *A reunião $F = F_1 \cup \cdots \cup F_k$ de um número finito de conjuntos fechados $F_1, \ldots, F_k$ é um conjunto fechado;*

3) *A interseção $F = \bigcap_{\lambda \in L} F_\lambda$ de uma família qualquer $(F_\lambda)_{\lambda \in L}$ de conjuntos fechados F_λ é um conjunto fechado.*

Demonstração: A afirmação 1) é evidente. Quanto a 2), se $F_1, \ldots, F_k$ são fechados então $A_1 = \complement F_1, \ldots, A_k = \complement F_k$ são abertos, portanto $A_1 \cap \cdots \cap A_k$ é aberto. Logo $F_1 \cup \cdots \cup F_k = \complement A_1 \cup \cdots \cup \complement A_k = \complement(A_1 \cap \cdots \cap A_k)$ é fechado. Finalmente, se cada F_λ, $\lambda \in L$, é fechado então cada $A_\lambda = \complement F_\lambda$ é aberto, logo $A = \bigcup_{\lambda \in L} A_\lambda$ também é aberto. Sendo assim, o conjunto $F = \bigcap_{\lambda \in L} F_\lambda = \bigcap_{\lambda \in L} \complement A_\lambda = \complement \left(\bigcup_{\lambda \in L} A_\lambda \right) = \complement A$ é fechado.

Note que uma reunião infinita de conjuntos fechados pode ser fechada ou não. De fato, para cada ponto $x \in \mathbb{R}^n$, o conjunto $\{x\}$ é fechado. Ora, todo conjunto $X \subset \mathbb{R}^n$ é reunião dos seus pontos: $X = \bigcup_{\lambda \in X} \{\lambda\}$. Como há conjuntos em $\mathbb{R}^n$ que não são fechados, há reuniões (infinitas) de conjuntos fechados que não são fechadas.

Segue-se da definição de fronteira que um ponto a pertence à fronteira do conjunto X se, e somente se, a é aderente a X e a $\mathbb{R}^n - X$. Ou seja, $\partial X = \overline{X} \cap (\overline{\mathbb{R}^n - X})$. Em particular, a fronteira de todo conjunto $X \subset \mathbb{R}^n$ é um conjunto fechado.

Fixemos um conjunto $X \subset \mathbb{R}^n$. Um subconjunto $F \subset X$ diz-se *fechado em* X quando se tem $F = X \cap G$, onde G é um conjunto fechado em $\mathbb{R}^n$. A fim de que o subconjunto $F \subset X$ seja fechado em X é necessário e suficiente que F contenha todos os seus pontos aderentes *que pertençam a* X.

Se $X \subset \mathbb{R}^n$ é fechado, então um subconjunto $F \subset X$ é fechado em X se, e somente se, é fechado em $\mathbb{R}^n$.

Os conjuntos fechados em X gozam de propriedades análogas às que foram demonstradas no Teorema 18 para os fechados em $\mathbb{R}^n$. A saber: $\varnothing$ e X são fechados em X; uma reunião finita ou uma interseção arbitrária de fechados em X é ainda um conjunto fechado em X.

Evidentemente, se $F \subset \mathbb{R}^n$ é fechado e $F \subset X$ então F é fechado em X. Seja $X = \{x \in \mathbb{R}; x > 0\}$ a semi-reta positiva aberta. O intervalo semi-aberto $(0, 1]$ é fechado em X.

(11.1) *Seja $F \subset X$. A fim de que F seja fechado em X é necessário e suficiente que o conjunto $A = X - F$ (complementar de F relativamente a X) seja aberto em X.*

Com efeito, dados $F' \subset \mathbb{R}^n$ e $A' = \complement F'$, o complementar de F' em $\mathbb{R}^n$, temos $F = X \cap F' \Leftrightarrow X - F = X \cap A'$. Ora, F' é fechado em $\mathbb{R}^n$ se, e somente se, A' é aberto. Assim, F é fechado em X se, e somente se $X - F$ é aberto em X.

A definição acima foi dada para que tivéssemos o

Teorema 19. *Seja $f\colon X \to \mathbb{R}^n$ uma aplicação definida no subconjunto $X \subset \mathbb{R}^m$. A fim de que f seja contínua, é necessário e suficiente que a imagem inversa $f^{-1}(F)$ de todo conjunto fechado $F \subset \mathbb{R}^n$ seja um conjunto fechado em X.*

Demonstração: Isto decorre do teorema análogo para abertos (Teorema 16) juntamente com o fato de que os conjuntos abertos (em X) são exatamente aqueles cujos complementares são fechados (em X). Com efeito, seja f contínua. Então, para cada $F \subset \mathbb{R}^n$ fechado, temos $F = \complement A$, onde $A \subset \mathbb{R}^n$ é aberto. Logo $f^{-1}(A)$ é aberto em X. Mas $f^{-1}(F) = f^{-1}(\complement A) = X - f^{-1}(A)$, portanto $f^{-1}(F)$ é fechado em X. Reciprocamente, se a imagem inversa por f de todo fechado em $\mathbb{R}^n$ é um fechado em X, então a relação $f^{-1}(A) = X - f^{-1}(F)$, com $A = \complement F$, mostra que a imagem inversa por f de todo aberto $A \subset \mathbb{R}^n$ é um aberto em X, portanto f é contínua.

Assim, por exemplo, se $f_1, \ldots, f_k\colon \mathbb{R}^n \to \mathbb{R}$ são funções reais contínuas então, dados os números reais $a_1, \ldots, a_k$, o conjunto X dos pontos $x \in \mathbb{R}^n$ para os quais se tem simultaneamente $f_1(x) \leq a_1, \ldots, f_k(x) \leq a_k$ é um subconjunto fechado de $\mathbb{R}^n$.

Em particular, fixado um ponto $a \in \mathbb{R}^n$, a função $f\colon \mathbb{R}^n \to \mathbb{R}$, dada por $f(x) = |x - a|$, é contínua. Logo, para todo $r > 0$, a bola fechada $B[a; r] = \{x \in \mathbb{R}^n; f(x) \leq r\}$ é um conjunto fechado.

Nas mesmas condições acima, é fechado o conjunto Y dos pontos $y \in \mathbb{R}^n$ para os quais se tem simultaneamente $f_1(y) = a_1, \ldots, f_k(y) = a_k$. Com efeito, vale $Y = f^{-1}(a)$, onde $f = (f_1, \ldots, f_k)\colon \mathbb{R}^n \to \mathbb{R}^k$ e $a = (a_1, \ldots, a_k)$.

Em particular, a esfera $S[a; r] = \{y \in \mathbb{R}^n; |y - a| = r\}$ é um subconjunto fechado de $\mathbb{R}^n$.

Observação: Se $f\colon X \to \mathbb{R}^n$ é contínua e $F \subset f(X)$ é um conjunto apenas fechado em $f(X)$, sua imagem inversa $f^{-1}(F)$ é ainda fechada

em X. Com efeito, temos $F = F' \cap f(X)$, onde F' é fechado em $\mathbb{R}^n$. Como $f^{-1}(F) = f^{-1}(F')$, segue-se do Teorema 19 que $f^{-1}(F)$ é fechada em X.

(11.2) *Se $F \subset \mathbb{R}^m$ e $G \subset \mathbb{R}^n$ são subconjuntos fechados então o produto cartesiano $F \times G \subset \mathbb{R}^m \times \mathbb{R}^n = \mathbb{R}^{m+n}$ é fechado.*

Com efeito, as projeções $\pi_1 \colon \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^m$ e $\pi_2 \colon \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^n$, definidas por $\pi_1(x,y) = x$ e $\pi_2(x,y) = y$, são contínuas e $F \times G = \pi_1^{-1}(F) \cap \pi_2^{-1}(G)$.

Um exemplo de conjunto fechado noutro é o gráfico de uma aplicação contínua $f \colon X \to \mathbb{R}^n$ definida num conjunto $X \subset \mathbb{R}^m$. Como sabemos, o gráfico de f é o subconjunto $G \subset X \times \mathbb{R}^n \subset \mathbb{R}^m \times \mathbb{R}^n$, formado pelos pontos $(x, f(x)) \in \mathbb{R}^m \times \mathbb{R}^n$, onde x varia em X. Afirmamos que G é um subconjunto fechado em $X \times \mathbb{R}^n$. (Quando $X \subset \mathbb{R}^m$ for fechado, então G será fechado no espaço euclidiano $\mathbb{R}^{m+n}$.) Com efeito, temos $G = \{(x,y) \in X \times \mathbb{R}^n; y - f(x) = 0\}$. Como a aplicação $\varphi \colon X \times \mathbb{R}^n \to \mathbb{R}^n$, dada por $\varphi(x,y) = y - f(x)$, é contínua, resulta que $G = \varphi^{-1}(0)$ é fechado em $X \times \mathbb{R}^n$.

Ao contrário do que ocorre para conjuntos abertos, a projeção $\pi_1 \colon \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^m$ não transforma necessariamente um conjunto fechado $F \subset \mathbb{R}^m \times \mathbb{R}^n$ num conjunto fechado $\pi_1(F) \subset \mathbb{R}^m$. Por exemplo, a hipérbole $H = \{(x,y) \in \mathbb{R}^2; xy = 1\}$ é um subconjunto fechado do plano, imagem inversa do fechado $\{1\} \subset \mathbb{R}$ pela aplicação contínua $(x,y) \mapsto xy$. Mas sua projeção no eixo das abcissas, $\pi_1(H) = \mathbb{R} - \{0\}$, não é fechada.

Dados $Y \subset X \subset \mathbb{R}^n$, podemos também definir o *fecho de Y relativamente a X* como sendo o conjunto $\overline{Y} \cap X$, dos pontos aderentes a Y que pertencem ao conjunto X. Então Y é fechado em X se, e somente se, coincide com seu fecho relativamente a X.

Um caso particular importante se dá quando o fecho de Y relativamente a X é todo o conjunto X. Para descrever esta situação, damos a definição abaixo.

Sejam $Y \subset X \subset \mathbb{R}^n$. Dizemos que Y é *denso em X* quando $\overline{Y} \cap X = X$, ou seja, $X \subset \overline{Y}$. Isto significa que todo ponto de X é limite de uma seqüência cujos termos pertencem a Y. Ou, ainda de outra maneira: dado $Y \subset X$, tem-se Y denso em X se, e somente se, toda bola aberta com centro em algum ponto de X contém pontos de Y.

Os fatos mais importantes a respeito de conjuntos densos em $\mathbb{R}^n$ são as seguintes observações.

(11.3) *Sejam* $f, g\colon X \to \mathbb{R}^n$ *aplicações contínuas num conjunto* $X \subset \mathbb{R}^m$ *e* $Y \subset X$ *um subconjunto denso. Se* $f(y) = g(y)$ *para todo* $y \in Y$ *então* $f = g$ *(isto é,* $f(x) = g(x)$ *para todo* $x \in X$*).*

Com efeito, para cada $x \in X$ existe uma seqüência de pontos $y_k \in Y$ com $\lim y_k = x$. Então $f(x) = f(\lim y_k) = \lim f(y_k) = \lim g(y_k) = g(\lim y_k) = g(x)$.

(11.4) *Todo conjunto* $X \subset \mathbb{R}^n$ *contém um subconjunto enumerável* E*, denso em* X*.*

Com efeito, a coleção $\mathcal{B}$ das bolas abertas $B(q;r)$, com centro num ponto $q \in \mathbb{Q}^n$ e raio racional, é enumerável: $\mathcal{B} = \{B_1, \ldots, B_i, \ldots\}$. Para cada $i \in \mathbb{N}$, escolhamos um ponto $x_i \in B_i \cap X$, se $B_i \cap X$ não for vazio. Caso seja $B_i \cap X = \varnothing$, x_i não existirá. O conjunto E dos pontos x_i assim obtidos é um subconjunto enumerável de X. Para mostrar que E é denso em X, tomemos arbitrariamente $x \in X$ e $\varepsilon > 0$. Existe $r > 0$ racional com $2r < \varepsilon$. Como $\mathbb{Q}^n$ é denso em $\mathbb{R}^n$, encontramos $q \in \mathbb{Q}^n$ tal que $|q - x| < r$. Logo $x \in B(q;r) = B_i$. Assim, $B_i \cap X \neq \varnothing$. Existe portanto $x_i \in E$. Como x e x_i pertencem à bola B_i, de raio r, temos $|x - x_i| < 2r < \varepsilon$. Ficou provado que toda bola aberta $B(x;\varepsilon)$ com centro em algum ponto de X contém um ponto $x_i \in E$. Então E é denso em X.

12 Conjuntos compactos

Diremos que um conjunto $K \subset \mathbb{R}^n$ é *compacto* quando ele for limitado e fechado.

Assim, por exemplo, são compactas todas as esferas e bolas fechadas do espaço euclidiano, mas o espaço $\mathbb{R}^n$ inteiro não é compacto (salvo se $n = 0$!).

Em virtude do Teorema de Bolzano-Weierstrass, um conjunto $K \subset \mathbb{R}^n$ é compacto se, e somente se, toda seqüência de pontos $x_k \in K$ possui uma subseqüência que converge para um ponto de K.

As seguintes propriedades decorrem imediatamente da definição:

(12.1) $K_1, \ldots, K_p$ compactos em $\mathbb{R}^n \Rightarrow K_1 \cup \cdots \cup K_p$ compacto.

(12.2) A interseção de uma família qualquer de compactos $K_\lambda \subset \mathbb{R}^n$ é um conjunto compacto.

(12.3) Se $K \subset \mathbb{R}^m$ e $L \subset \mathbb{R}^n$ são compactos então o produto cartesiano $K \times L \subset \mathbb{R}^{m+n}$ é compacto.

Menos óbvio é o seguinte fato, conhecido como a *propriedade de Cantor*:

(12.4) *Dada uma seqüência decrescente* $K_1 \supset \cdots \supset K_k \supset \ldots$ *de compactos não vazios, a interseção* $K = \bigcap_{k=1}^{\infty} K_k$ *(é compacta e) não é vazia.*

Demonstração: Escolhamos, para cada $k \in \mathbb{N}$, um ponto $x_k \in K_k$, obtendo assim uma seqüência, a qual possui uma subseqüência (x_{k_i}), $i \in \mathbb{N}$, convergindo para um ponto $x \in \mathbb{R}^n$. Dado arbitrariamente $k \in \mathbb{N}$, temos $x_{k_i} \in K_k$ para todo $k_i > k$, logo $x = \lim x_{k_i} \in K_k$. Assim, o ponto x pertence a K_k para todo $k \in \mathbb{N}$, ou seja $x \in K = \bigcap_k K_k$, o que mostra que K não é vazio.

Teorema 20. *Seja* $f\colon X \to \mathbb{R}^n$ *contínua no conjunto* $X \subset \mathbb{R}^m$. *Para todo subconjunto compacto* $K \subset X$*, sua imagem* $f(K)$ *é compacta.*

Demonstração: Mostremos primeiro que $f(K)$ é fechado em $\mathbb{R}^n$. Seja pois $y \in \mathbb{R}^n$ aderente a $f(K)$. Então $y = \lim f(x_k)$, $x_k \in K$ para todo $k \in \mathbb{N}$. Pela compacidade de K, uma subseqüência (x_{k_i}) converge para um certo ponto $x \in K$. Segue-se que $y = \lim_{i\to\infty} f(x_{k_i}) = f(\lim x_{k_i}) = f(x)$, donde $y \in f(K)$. Agora, mostremos que $f(K)$ é limitado. De fato, se não fosse, poderíamos, para cada $k \in \mathbb{N}$, obter um ponto $x_k \in K$ tal que $|f(x_k)| > k$. Então a seqüência $(f(x_k))$ não admitiria subseqüências convergentes. Mas (x_k) tem uma subseqüência convergente, com $\lim_{i\to\infty} x_{k_i} = x \in K$. A continuidade de f nos dá então $f(x) = f(\lim x_{k_i}) = \lim f(x_{k_i})$, uma contradição.

Note que uma aplicação contínua pode transformar um conjunto limitado num conjunto ilimitado $\big(f(x) = 1/x$ leva o intervalo aberto $(0,1)$ no intervalo infinito $(1,+\infty)\big)$ ou um conjunto fechado num conjunto não fechado $\big(f(x) = 1/(1+x^2)$ leva $\mathbb{R}$ no intervalo $(0,1]\big)$. Mas se $X \subset \mathbb{R}^m$ for limitado e fechado, sua imagem por qualquer aplicação contínua é limitada e fechada.

Corolário 1 (Weierstrass.) *Toda função real contínua* $f\colon K \to \mathbb{R}$*, definida num compacto* $K \subset \mathbb{R}^m$*, atinge seu máximo e seu mínimo em* K*, isto é, existem pontos* $x_0, x_1 \in K$ *tais que* $f(x_0) \le f(x) \le f(x_1)$ *para qualquer* $x \in K$.

Com efeito $f(K) \subset \mathbb{R}$ é compacto, logo $y_0 = \inf f(K)$ e $y_1 = \sup f(K)$ pertencem a $f(K)$, isto é, existem pontos $x_0, x_1 \in K$ tais

que $f(x_0) = y_0$ e $f(x_1) = y_1$. Então $f(x_0) \leq f(x) \leq f(x_1)$ para todo $x \in K$.

A função $f\colon \mathbb{R} \to \mathbb{R}$, definida por $f(x) = x/(1+|x|)$, é contínua, cumpre $-1 < f(x) < 1$ para todo $x \in \mathbb{R}$, mas nenhum valor $f(x)$ é menor nem maior do que todos os demais valores de f.

Em particular, toda aplicação contínua $f\colon K \to \mathbb{R}^n$, definida num compacto $K \subset \mathbb{R}^m$, é limitada, isto é, existe $c > 0$ tal que $|f(x)| \leq c$ para todo $x \in K$.

Note ainda que se $f\colon K \to \mathbb{R}$ é contínua no compacto $K \subset \mathbb{R}^n$ e $f(x) > 0$ para todo $x \in K$ então existe $\varepsilon > 0$ tal que $f(x) \geq \varepsilon$ para todo $x \in K$. Isto seria falso se K não fosse compacto; bastaria tomar $f\colon (0, +\infty) \to \mathbb{R}$, dada por $f(x) = 1/x$.

Corolário 2. *Seja $K \subset \mathbb{R}^m$ compacto. Toda aplicação contínua $f\colon K \to \mathbb{R}^n$ é fechada, isto é, $F \subset K$ fechado $\Rightarrow f(F) \subset \mathbb{R}^n$ fechado.*

Com efeito, $F \subset K$ fechado $\Rightarrow F$ compacto $\Rightarrow f(F)$ compacto $\Rightarrow f(F)$ fechado em $\mathbb{R}^n$.

Corolário 3. *A inversa de uma bijeção contínua definida num compacto é contínua. (Ou seja: toda bijeção contínua definida num compacto é um homeomorfismo sobre sua imagem.)*

Com efeito, seja $f\colon K \to L$ uma bijeção contínua do compacto $K \subset \mathbb{R}^m$ no compacto $L \subset \mathbb{R}^n$. Pondo $g = f^{-1}\colon L \to K$ vemos que, para todo $F \subset K$ fechado, que imagem inversa $g^{-1}(F) = f(F)$ é fechada em $\mathbb{R}^n$, pelo Corolário 2. Segue-se do Teorema 19 que g é contínua.

Vemos agora que só foi possível definir a bijeção contínua $f\colon [0, 2\pi) \to S^1$ com inversa descontínua porque seu domínio é um intervalo não-compacto.

Uma conseqüência do Corolário 2 é a seguinte:

(12.5) *Seja $\varphi\colon K \to L$ uma aplicação contínua do compacto $K \subset \mathbb{R}^m$ sobre o conjunto (necessariamente compacto) $L = \varphi(K) \subset \mathbb{R}^n$. Dado $F \subset L$, se sua imagem inversa $\varphi^{-1}(F)$ é fechada, então F é fechado.*

Com efeito, como φ é sobrejetiva, $\varphi[\varphi^{-1}(F)] = F$. Pelo Corolário 2, F é fechado. Este fato tem como corolário:

(12.6) *Seja $\varphi\colon K \to L$ uma aplicação contínua do compacto $K \subset \mathbb{R}^m$ sobre o compacto $L \subset \mathbb{R}^n$. Então uma aplicação $f\colon L \to \mathbb{R}^p$ é contínua se, e somente se, $f \circ \varphi\colon K \to \mathbb{R}^p$ é contínua.*

Se f é contínua, evidentemente $f \circ \varphi$ também é. Reciprocamente, supondo $f \circ \varphi$ contínua, para cada $F \subset \mathbb{R}^p$ fechado temos $\varphi^{-1}[f^{-1}(F)] =$

$(f \circ \varphi)^{-1}(F)$ fechado em K. Pelo que vimos acima, isto implica $f^{-1}(F)$ fechado em L. Logo f é contínua (Teorema 19).

Como aplicação, seja $g\colon [0, 2\pi] \to \mathbb{R}^n$ contínua tal que $g(0) = g(2\pi)$. Através de g, podemos definir uma aplicação contínua $f\colon S^1 \to \mathbb{R}^n$, pondo $f(e^{it}) = f(\cos t, \operatorname{sen} t) = g(t)$, $0 \leq t \leq 2\pi$. Como $g(0) = g(2\pi)$, f está bem definida. Seja $\varphi\colon [0, 2\pi] \to S^1$ a sobrejeção contínua dada por $\varphi(t) = e^{it} = (\cos t, \operatorname{sen} t)$. Então $f \circ \varphi = g$ é contínua. Pelo resultado acima, $f\colon S^1 \to \mathbb{R}^n$ é contínua. Isto se exprime dizendo que "definir uma aplicação contínua no círculo S^1 é o mesmo que definí-la no intervalo $[0, 2\pi]$, assumindo valores iguais nos extremos".

Teorema 21. *Toda aplicação contínua $f\colon K \to \mathbb{R}^n$, definida num compacto $K \subset \mathbb{R}^m$, é uniformemente contínua.*

Às vezes é necessário usar a seguinte versão, mais abrangente, do teorema acima:

Teorema 21a. *Se $f\colon X \to \mathbb{R}^n$ é contínua e $K \subset X$ é compacto então, para todo $\varepsilon > 0$, existe $\delta > 0$ tal que $x \in X$, $y \in K$, $|x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon$.*

Demonstração: Suponha, por absurdo, que o teorema fosse falso. Então existiriam $\varepsilon > 0$ e duas seqüências de pontos $x_k \in X$, $y_k \in K$ tais que $|x_k - y_k| < 1/k$ e $|f(x_k) - f(y_k)| \geq \varepsilon$, para todo $k \in \mathbb{N}$. Passando a uma subseqüência, se necessário, podemos supor que $\lim y_k = y \in K$, donde $\lim x_k = y$ também. Então, pela continuidade de f, viria $\varepsilon \leq \lim |f(x_k) - f(y_k)| = |f(y) - f(y)|$ uma contradição.

Outro tipo de uniformidade ocorre quando uma função (contínua) depende continuamente de um parâmetro α que varia num compacto. Mais precisamente:

Teorema 21b. *Seja $f\colon X \times K \to \mathbb{R}^n$ contínua, onde K é compacto. Fixemos $x_0 \in X$. Para todo $\varepsilon > 0$, existe $\delta > 0$ tal que $x \in X$, $|x - x_0| < \delta \Rightarrow |f(x, \alpha) - f(x_0, \alpha)| < \varepsilon$, seja qual for $\alpha \in K$.*

Demonstração: Supondo o contrário, existiriam $\varepsilon > 0$ e seqüências de pontos $x_k \in X$, $\alpha_k \in K$ tais que $|x_k - x_0| < 1/k$ e $|f(x_k, \alpha_k) - f(x_0, \alpha_k)| \geq \varepsilon$. Passando a uma subseqüência, se necessário, podemos admitir que $\lim \alpha_k = \alpha \in K$. Como, evidentemente, $\lim x_k = x_0$, a continuidade de f nos daria $\varepsilon \leq \lim |f(x_k, \alpha_k) - f(x_0, \alpha_k)| = |f(x_0, \alpha) - f(x_0, \alpha)|$, uma contradição.

Uma aplicação do teorema acima é:

(12.7) *Seja* $f\colon X\times[a,b]\to\mathbb{R}$ *contínua. Definamos* $\varphi\colon X\to\mathbb{R}$ *pondo, para cada* $x\in X$,

$$\varphi(x)=\int_a^b f(x,t)\,dt.$$

Então φ *é contínua em cada ponto* $x_0\in X$.

Com efeito, $|\varphi(x)-\varphi(x_0)|\le\int_a^b|f(x,t)-f(x_0,t)|dt$. Pelo Teorema 21b, dado $\varepsilon>0$, podemos achar $\delta>0$ tal que $x\in X$, $|x-x_0|<\delta\Rightarrow|f(x,t)-f(x_0,t)|<\varepsilon/(b-a)$, seja qual for $t\in[a,b]$, logo tem-se $|\varphi(x)-\varphi(x_0)|<\varepsilon$.

Caracterizaremos agora a compacidade por meio de coberturas.

Uma *cobertura* de um conjunto $X\subset R^n$ é uma família $(C_\lambda)_{\lambda\in L}$ de subconjuntos $C_\lambda\subset\mathbb{R}^n$ tal que $X\subset\bigcup_{\lambda\in L}C_\lambda$. Isto significa que, para cada $x\in X$, existe um $\lambda\in L$ tal que $x\in C_\lambda$.

Uma *subcobertura* é uma subfamília $(C_\lambda)_{\lambda\in L'}$, $L'\subset L$, tal que ainda se tem $X\subset\bigcup_{\lambda\in L'}C_\lambda$.

Diz-se que a cobertura $X\subset\cup C_\lambda$ é *aberta* quando os C_λ forem todos abertos, *finita* se L é um conjunto finito, *enumerável* se L é enumerável, etc.

Teorema 22 (Lindelöf). *Seja* $X\subset\mathbb{R}^n$ *um conjunto arbitrário. Toda cobertura aberta* $X\subset\cup A_\lambda$ *admite uma subcobertura enumerável* $X\subset A_{\lambda_1}\cup\cdots\cup A_{\lambda_i}\cup\dots$

Demonstração: Seja $E=\{x_1,\dots,x_i,\dots\}\subset X$ um subconjunto enumerável, denso em X. Consideremos o conjunto $\mathcal{B}$ de todas as bolas abertas $B(x;r)$, com centro num ponto de E, raio racional e tais que cada uma delas está contida em algum $A_\lambda\cdot\mathcal{B}$ é um conjunto enumerável de bolas. Afirmamos que as bolas $B\in\mathcal{B}$ cobrem X. Com efeito, dado $x\in X$, existe $\lambda\in L$ tal que $x\in A_\lambda$. Como A_λ é aberto, existe $r>0$ racional tal que $B(x;2r)\subset A_\lambda$. Sendo E denso em X, podemos encontrar $x_i\in E$ com $|x-x_i|<r$. Então $x\in B(x_i;r)$. Para mostrar que $B(x_i;r)\in\mathcal{B}$, resta ver que esta bola está contida em A_λ. Ora, $y\in B(x_i;r)\Rightarrow|y-x_i|<r\Rightarrow|y-x|\le|y-x_i|+|x_i-x|<2r\Rightarrow y\in B(x;2r)\subset A_\lambda$. Isto conclui a verificação de que as bolas $B\in\mathcal{B}$ cobrem X. Tomando uma enumeração $B_1,\dots,B_i,\dots$ para essas bolas e escolhendo, para cada $i\in\mathbb{N}$, um índice $\lambda_i\in L$ tal que $B_i\subset A_{\lambda_i}$, concluimos que $X\subset A_{\lambda_1}\cup\cdots\cup A_{\lambda_i}\cup\dots$

Teorema 23 (Borel-Lebesgue). *Seja* $K\subset\mathbb{R}^n$ *compacto (isto é, limitado*

e fechado). Toda cobertura aberta $K \subset \bigcup_{\lambda \in L} A_\lambda$ *admite uma subcobertura finita* $K \subset A_{\lambda_1} \cup \cdots \cup A_{\lambda_i}$.

Demonstração: Pelo Teorema de Lindelöf, obtemos uma subcobertura enumerável $K \subset A_{\lambda_1} \cup \cdots \cup A_{\lambda_i} \cup \ldots$. Ponhamos $K_i = K \cap \mathcal{C}(A_{\lambda_1} \cup \cdots \cup A_{\lambda_i})$ para cada $i \in \mathbb{N}$. Isto nos dá uma seqüência decrescente $K_1 \supset K_2 \supset \cdots \supset K_i \supset \ldots$ de compactos. Dado qualquer $x \in K$, existe algum $i \in \mathbb{N}$ tal que $x \in A_{\lambda_i}$. Então $x \notin K_i$. Isto mostra que nenhum ponto $x \in K$ está em todos os K_i, ou seja, que $\bigcap_{i=1}^{\infty} K_i = \varnothing$. Segue-se então da propriedade de Cantor que algum dos compactos K_i é vazio, o que significa $K \subset A_{\lambda_1} \cup \cdots \cup A_{\lambda_i}$.

Vale também a recíproca do Teorema de Borel-Lebesgue:

Teorema 24. *Se toda cobertura aberta do conjunto* $K \subset \mathbb{R}^m$ *admite uma subcobertura finita, então* K *é limitado e fechado (isto é, compacto).*

Demonstração: Em primeiro lugar, as bolas abertas de raio 1 e centros nos pontos de K constituem uma cobertura aberta $K \subset \bigcup_{x \in K} B(x;1)$, a qual possui uma subcobertura finita $K \subset B(x_1;1) \cup \cdots \cup B(x_i,1)$. Assim, K está contido numa reunião finita de conjuntos limitados, logo é limitado. Além disso, K é fechado pois, se não fosse, existiria um ponto $a \in \overline{K} - K$. Então, para cada $i \in \mathbb{N}$, tomamos $A_i =$ complementar da bola fechada $B[a;1/i]$. Para todo $x \in K$, temos $x \neq a$, logo $|x-a| > 1/i$ para algum i, o que nos dá $x \in A_i$. Portanto $K \subset \bigcup_{i=1}^{\infty} A_i$, uma cobertura aberta, da qual extraímos uma subcobertura finita: $K \subset A_{i_1} \cup \cdots \cup A_{i_p}$. Como $A_1 \subset A_2 \subset \cdots \subset A_i \subset \ldots$, toda reunião de uma coleção finita de conjuntos A_i é igual ao conjunto de maior índice na coleção. Assim, temos $K \subset A_i$ para algum i. Esta inclusão significa que a bola $B[a;1/i]$ não tem pontos em comum com K, o que contradiz ser $a \in \overline{K}$ e prova o teorema.

Os Teoremas 23 e 24 mostram que poderíamos, equivalentemente, ter definido um conjunto compacto K pela condição de que toda cobertura aberta $K \subset \cup A_\lambda$ admita uma subcobertura finita $K \subset A_{\lambda_1} \cup \cdots \cup A_{\lambda_i}$. Tal definição é, de fato, a mais conveniente para estudos mais gerais. A que demos no texto é mais simples, porém interessante apenas para espaços euclidianos.

Como aplicação do Teorema de Borel-Lebesgue, demonstraremos o seguinte:

(12.8) *Se o aberto* U *contém a interseção* $K = \bigcap_{i=1}^{\infty} K_i$ *de uma seqüência decrescente* $K_1 \supset K_2 \supset \cdots \supset K_i \supset \ldots$ *de conjuntos compactos, então existe* $i \in \mathbb{N}$ *tal que* $K_i \subset U$.

Com efeito, os abertos $U_i = \mathbb{R}^n - K_i$, juntamente com U, constituem uma cobertura de K_1, da qual extraimos uma subcobertura finita $K_1 \subset U \cup U_{i_1} \cup \cdots \cup U_{i_p}$. Seja i o maior dos índices $i_1, \ldots, i_p$. Como $U_1 \subset U_2 \subset \ldots$, temos $U_{i_1} \cup \cdots \cup U_{i_p} = U_i$, logo $K_1 \subset U \cup U_i$. Com maior razão, $K_i \subset U \cup U_i$. Como nenhum ponto de K_i pode pertencer ao seu complementar U_i, devemos ter $K_i \subset U$, como queríamos demonstrar.

13 Distância entre dois conjuntos; diâmetro

Sejam $S, T \subset \mathbb{R}^n$ conjuntos não vazios. Definiremos a *distância* $d(S,T)$ entre S e T pondo

$$d(S,T) = \inf\{|x-y|; x \in S,\ y \in T\}.$$

Evidentemente, $d(S,T) = d(T,S)$, $S \cap T \neq \varnothing \Rightarrow d(S,T) = 0$ e $S_1 \subset S_2$, $T_1 \subset T_2 \Rightarrow d(S_2,T_2) \leq d(S_1,T_1)$.

De acordo com a definição de ínfimo de um conjunto de números reais, a distância $d(S,T)$ é caracterizada pelas propriedades abaixo:

1) Tem-se $d(S,T) \leq |x-y|$ para quaisquer $x \in S$, $y \in T$;

2) Dado arbitrariamente $\varepsilon > 0$, existem $x \in S$ e $y \in T$ tais que $|x-y| < d(S,T) + \varepsilon$.

Por exemplo, se $S, T \subset \mathbb{R}^3$ são duas retas não contidas no mesmo plano, então $d(S,T)$ é o comprimento do (único) segmento de reta perpendicular a S e a T, ligando essas retas.

Um caso particular importante da distância entre dois conjuntos ocorre quando um deles se reduz a um ponto.

Dados $x \in \mathbb{R}^n$ e um conjunto não vazio $T \subset \mathbb{R}^n$, temos:

$$d(x,T) = \inf\{|x-y|; y \in T\}.$$

Evidentemente, $x \in T \Rightarrow d(x,T) = 0$ e $T_1 \subset T_2 \Rightarrow d(x,T_2) \leq d(x,T_1)$.

Novamente valem as propriedades características:

1) Tem-se $d(x,T) \leq |x-y|$, qualquer que seja $y \in T$;

2) Dado arbitrariamente $\varepsilon > 0$, existe $y \in T$ tal que $|x-y| < d(x,T)+\varepsilon$.

Por exemplo, se $x \in \mathbb{R}^2$ e $T \subset \mathbb{R}^2$ é uma reta que não contém x, então $d(x,T)$ é o comprimento do segmento de perpendicular baixado de x sobre T.

O ínfimo de um conjunto de números não-negativos é igual a zero se, e somente se, tal conjunto contém números arbitrariamente pequenos. Assim, tem-se $d(x,T)=0$ se, e somente se, dado qualquer $\varepsilon > 0$, existe $y \in T$ tal que $d(x,y) < \varepsilon$. Em outras palavras,

$$d(x,T) = 0 \Leftrightarrow x \in \overline{T}.$$

Em particular, se $F \subset \mathbb{R}^n$ é um conjunto fechado, vale $d(x,F) = 0 \Leftrightarrow x \in F$.

Como $\partial T = \overline{T} \cap \overline{\mathbb{R}^n - T}$, vemos que $x \in \partial T$ se, e somente se, $d(x,T) = d(x, \mathbb{R}^n - T) = 0$.

Sempre que usarmos os símbolos $d(S,T)$ e $d(x,T)$, estaremos admitindo tacitamente que $S \neq \varnothing$ e $T \neq \varnothing$.

Teorema 25. $d(S,T) = d(\overline{S},\overline{T})$.

Demonstração: Evidentemente, $d(\overline{S},\overline{T}) \leq d(S,T)$. Para provar a desigualdade oposta, seja ε um número positivo arbitrário. Existem $\bar{x} \in \overline{S}$ e $\bar{y} \in \overline{T}$ tais que $|\bar{x}-\bar{y}| < d(\overline{S},\overline{T}) + \varepsilon/3$, e $x \in S$, $y \in T$ tais que $|x-\bar{x}| < \varepsilon/3$, $|y-\bar{y}| < \varepsilon/3$. Então

$$d(S,T) \leq |x-y| \leq |x-\bar{x}| + |\bar{x}-\bar{y}| + |\bar{y}-y| < \varepsilon/3 + \\ + d(\overline{S},\overline{T}) + \varepsilon/3 + \varepsilon/3.$$

Logo $d(S,T) < d(\overline{S},\overline{T}) + \varepsilon$. Como $\varepsilon > 0$ é arbitrário, concluimos que $d(S,T) \leq d(\overline{S},\overline{T})$, o que prova o teorema.

Corolário. $d(x,T) = d(x,\overline{T})$.

Teorema 26. *Se $K \subset \mathbb{R}^n$ é compacto e $F \subset \mathbb{R}^n$ é fechado, então existem $x_0 \in K$ e $y_0 \in F$ tais que $d(K,F) = |x_0 - y_0|$. Em particular, se $K \cap F = \varnothing$ então $d(K,F) > 0$.*

Demonstração: Como o ínfimo de um conjunto de números reais é limite de uma seqüência de elementos desse conjunto, temos $d(K,F) =$

$\lim_{k\to\infty} |x_k - y_k|$ para seqüências convenientes de pontos $x_k \in K$ e $y_k \in F$. Da desigualdade $|y_k| \le |y_k - x_k| + |x_k|$ resulta que a seqüência (y_k) é limitada, pois são limitadas $(|x_k - y_k|)$, por ser convergente, e (x_k), por pertencerem seus termos x_k ao compacto K. Passando a subseqüências, se necessário, podemos admitir, portanto, que $x_k \to x_0$ e $y_k \to y_0$. Como K e F são fechados, temos $x_0 \in K$ e $y_0 \in F$. Além disso, $d(K, F) = \lim_{k\to\infty} |x_k - y_k| = |x_0 - y_0|$.

Corolário 1. *Seja $x \in \mathbb{R}^n$. Se $F \subset \mathbb{R}^n$ é fechado, então existe $y_0 \in F$ tal que $d(x, F) = |x - y_0|$.*

Corolário 2. *Sejam $K \subset \mathbb{R}^n$ compacto e $U \subset \mathbb{R}^n$ aberto. Se $K \subset U$ então existe $\varepsilon > 0$ tal que $x \in K \Rightarrow B(x;\varepsilon) \subset U$. Em particular, $x \in K$, $y \in \mathbb{R}^n$, $|x - y| < \varepsilon \Rightarrow [x, y] \subset U$.*

Com efeito, $F = \mathbb{R}^n - U$ é fechado e $K \cap F = \varnothing$. Seja $d(K, F) = \varepsilon > 0$. Se $y \in F$ e $x \in K$ então $\varepsilon = d(K, F) \le |x - y|$, logo $y \notin B(x;\varepsilon)$. Assim, para este valor fixo de ε, toda bola aberta de raio ε com centro num ponto de K está contida em U.

Corolário 3. *Sejam $S, T \subset \mathbb{R}^n$, com S limitado. Existem $x_0 \in \overline{S}$, $y_0 \in \overline{T}$ tais que $|x_0 - y_0| = d(S, T)$.*

Com efeito, $d(S, T) = d(\overline{S}, \overline{T})$, onde $\overline{S}$ é compacto e $\overline{T}$ é fechado.

(13.1) **Exemplo**. Em geral, existem muitos pontos do conjunto fechado F que estão a uma distância mínima do ponto x. Um caso extremo disso é $F = S[x; r]$: para todo $y \in F$, tem-se $|x - y| = d(x, F)$! Mas, quando $F \subset \mathbb{R}^n$ é convexo e a norma de $\mathbb{R}^n$ provém de um produto interno, prova-se que existe um *único* $y_0 \in F$ tal que $d(x, F) = |x - y_0|$. (Com efeito, se $x_0 \ne y_0 \in F$ são tais que $|x_0 - x| = |y_0 - x|$ então $z_0 = \dfrac{x_0 + y_0}{2} \in F$ é o ponto médio da base do triângulo isósceles xx_0y_0, logo $|z_0 - x| < |x_0 - x|$.)

(13.2) **Exemplo**. Dados dois conjuntos fechados ilimitados $F, G \subset \mathbb{R}^n$, pode-se ter $F \cap G = \varnothing$ com $d(F, G) = 0$. Basta tomar o eixo das abcissas $F = \{(x, 0) \in \mathbb{R}^2; x \in \mathbb{R}\}$ e o ramo de hipérbole $G = \{(x, 1/x) \in \mathbb{R}^2; x > 0\}$. Neste caso, não existem $x_0 \in F$, $y_0 \in G$ tais que $|x_0 - y_0| = d(F, G)$.

Teorema 27. $|d(x, T) - d(y, T)| \le |x - y|$.

Demonstração: Existem $x_0, y_0 \in \overline{T}$ tais que $d(x, T) = d(x, \overline{T}) = |x -$

$x_0|$ e $d(y,T) = d(y,\overline{T}) = |y - y_0|$. Então $|x - x_0| \leq |x - y_0|$. Daí

$$d(x,T) = |x - x_0| \leq |x - y_0| \leq |x - y| + +|y - y_0| \\ = |x - y| + d(y,T),$$

isto é, $d(x,T) - d(y,T) \leq |x - y|$. Analogamente, $d(y,T) - d(x,T) \leq |x - y|$. Logo $|d(x,T) - d(y,T)| \leq |x - y|$.

Corolário. *A função $f\colon \mathbb{R}^n \to \mathbb{R}$, definida por $f(x) = d(x,T)$, é uniformemente contínua.*

Com efeito, pelo Teorema 27, f é uma contração fraca.

(13.3) **Exemplo**. Sejam $F, G \subset \mathbb{R}^n$ dois subconjuntos fechados disjuntos e não-vazios. A função $f\colon \mathbb{R}^n \to \mathbb{R}$, definida por $f(x) = \dfrac{d(x,F)}{d(x,F) + d(x,G)}$ é contínua, assume o valor 0 em todos os pontos de F e o valor 1 em todos os pontos de G. Ela se chama uma *função de Urysohn* do par (F,G). Pondo $A = f^{-1}((-\infty, 1/2))$ e $B = f^{-1}((1/2, +\infty))$ vemos que A e B são abertos disjuntos, com $F \subset A$ e $G \subset B$. Concluimos então que, dados dois fechados disjuntos F, G em $\mathbb{R}^n$, existem sempre abertos disjuntos A, B tais que $F \subset A$ e $G \subset B$.

Na mesma ordem das idéias acima acha-se a noção de *diâmetro* de um conjunto limitado não vazio $T \subset \mathbb{R}^n$. Definimos

$$\operatorname{diam}(T) = \sup\{|x - y|; x, y \in T\}.$$

Para quaisquer $x, y \in T$, tem-se $|x - y| \leq \operatorname{diam}(T)$ e, dado arbitrariamente $\varepsilon > 0$ existem $x, y \in T$ tais que $\operatorname{diam}(T) - \varepsilon < |x - y|$.

(13.4) Existem seqüências de pontos $x_k, y_k \in T$ tais que $\lim\limits_{k\to\infty} |x_k - y_k| = \operatorname{diam}(T)$. Como T é limitado, passando a subseqüências se necessário, podemos supor que existem $\lim x_k = x_0$ e $\lim y_k = y_0$. Então $x_0, y_0 \in \overline{T}$ e $\lim |x_k - y_k| = |x_0 - y_0|$, logo $|x_0 - y_0| = \operatorname{diam}(T)$. Quando T é fechado (donde compacto) temos $x_0, y_0 \in T$, logo o diâmetro de um conjunto compacto é a maior distância entre dois dos seus pontos.

Se $S \subset T$ então $\operatorname{diam}(S) \leq \operatorname{diam}(T)$.

(13.5) *O diâmetro da bola fechada $B[a;r]$* é igual a $2r$.

Com efeito, $x, y \in B]a;r] \Rightarrow |x - a| \leq r, \quad |y - a| \leq r \Rightarrow |x - y| \leq |x-a|+|y-a| \leq 2r$, logo $\operatorname{diam} B[a;r] \leq 2r$. Por outro lado, se tomarmos um vetor $v \in \mathbb{R}^n$ com $|v| = r$ e pusermos $x = a + v, \quad y = a - v$ então $x, y \in B[a;r]$ e $|x - y| = |2v| = 2r$, logo $\operatorname{diam} B[a;r] = 2r$.

Se $T \subset B[a;r]$ então $\operatorname{diam}(T) \leq 2r$.

Por outro lado, se $\operatorname{diam}(T) = r$ então, dado qualquer ponto $a \in T$ tem-se $|x - a| \leq r$, seja qual for $x \in T$. Portanto $T \subset B[a; r]$.

Uma seqüência do teorema abaixo é que também o diâmetro da bola aberta $B(a; r)$ é igual a $2r$.

Teorema 28. *Para todo conjunto limitado não-vazio $T \subset \mathbb{R}^n$, vale* $\operatorname{diam}(\overline{T}) = \operatorname{diam}(T)$.

Demonstração: Como $T \subset \overline{T}$, vale $\operatorname{diam}(T) \leq \operatorname{diam}(\overline{T})$. Por outro lado, dado arbitrariamente $\varepsilon > 0$, existem pontos $\bar{x}, \bar{y} \in \overline{T}$ tais que $\operatorname{diam}(\overline{T}) - \varepsilon/3 < |\bar{x} - \bar{y}|$ e $x, y \in T$ tais que $|x - \bar{x}| < \varepsilon/3$, $|y - \bar{y}| < \varepsilon/3$. Então $\operatorname{diam}(\overline{T}) - \varepsilon/3 \leq |\bar{x} - \bar{y}| \leq |\bar{x} - x| + |x - y| + |y - \bar{y}| < \operatorname{diam}(T) + 2\varepsilon/3$, donde $\operatorname{diam}(\overline{T}) \leq \operatorname{diam}(T) + \varepsilon$. Como $\varepsilon > 0$ é qualquer, vemos que $\operatorname{diam}(\overline{T}) \leq \operatorname{diam}(T)$.

Teorema 29. *Seja $f: K \to U$ uma aplicação contínua, definida num compacto $K \subset \mathbb{R}^m$ e tomando valores num aberto $U \subset \mathbb{R}^n$. Existe $\delta > 0$ tal que a imagem $f(T)$ de qualquer subconjunto $T \subset K$ com* $\operatorname{diam}(T) < \delta$ *está contida em alguma bola $B \subset U$.*

Demonstração: A imagem $f(K)$ é um compacto contido no aberto U. Pelo Corolário 2 do Teorema 26, existe $\varepsilon > 0$ tal que, para todo $x \in K$, a bola $B(f(x); \varepsilon)$ está contida em U. Pela continuidade uniforme de f, existe $\delta > 0$ tal que $x, y \in K, \quad |x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon$. Seja $T \subset K$ um subconjunto (não-vazio) com $\operatorname{diam}(T) < \delta$. Tomemos $x_0 \in T$. Então $x \in T \Rightarrow |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon \Rightarrow f(x) \in B(f(x_0); \varepsilon) = B$. Logo $f(T) \subset B \subset U$.

Observação: Ao enunciado do Teorema 29 podemos, se for conveniente, acrescentar que as bolas B podem ser todas tomadas com o mesmo raio ε.

Nesta ordem de idéias, existe um conceito que freqüentemente se mostra útil em demonstrações envolvendo conjuntos compactos. Trata-se da noção de número de Lebesgue.

Diz-se que um número $\delta > 0$ é *número de Lebesgue* de uma cobertura $X \subset \bigcup_{\lambda \in L} C_\lambda$ quando todo subconjunto de X com diâmetro menor do que δ está inteiramente contido em algum C_λ.

Uma cobertura, mesmo aberta e finita, pode não ter número de Lebesgue algum. Por exemplo, seja $\mathbb{R} - \{0\} = (-\infty, 0) \cup (0, +\infty)$ a cobertura de $\mathbb{R} - \{0\}$ pelas duas semi-retas $(-\infty, 0)$ e $(0, +\infty)$. Dado qualquer $\delta > 0$ podemos tomar um número a tal que $0 < a < \delta/2$.

Então o conjunto $A = \{-a, a\}$ tem diâmetro menor do que δ mas não está contido em nenhum dos dois conjuntos da cobertura dada. Logo, não existe número de Lebesgue para tal cobertura.

Vale entretanto o resultado seguinte:

(13.6) *Toda cobertura aberta $X \subset \bigcup A_\lambda$ de um conjunto compacto $X \subset \mathbb{R}^n$ possui número de Lebesgue.*

Demonstração. Se supusermos, por absurdo, que nenhum $\delta > 0$ é número de Lebesgue da cobertura dada, obteremos, para cada $k \in \mathbb{N}$, um conjunto $Y_k \subset X$, com diam $Y_k < 1/k$ mas nenhum A_λ contém Y_k. Escolhemos um ponto y_k em cada Y_k. Passando a uma subseqüência, se necessário, a compacidade de X assegura a existência de $a \in X$ tal que $\lim y_k = a$. Existe λ_0 tal que $a \in A_{\lambda_0}$. Existe ainda $\varepsilon > 0$ com $B(a;\varepsilon) \subset A_{\lambda_0}$, pois A_{λ_0} é aberto. Tomemos $k \in \mathbb{N}$ tão grande que $1/k < \varepsilon/2$ e $|a - y_k| < \varepsilon/2$. Então, lembrando que diam $Y_k < 1/k$, para todo $y \in Y_k$ temos $|y - a| \leq |y - y_k| + |y_k - a| < (1/k) + (\varepsilon/2) < \varepsilon$. Segue-se que $Y_k \subset B(a, \varepsilon) \subset A_{\lambda_0}$, uma contradição. □

14 Conexidade

Uma *cisão* de um subconjunto $X \subset \mathbb{R}^n$ é uma decomposição $X = A \cup B$, onde $A \cap B = \varnothing$ e os conjuntos A, B são ambos abertos em X.

Todo conjunto $X \subset \mathbb{R}^n$ admite pelo menos a *cisão trivial* $X = X \cup \varnothing$. Um exemplo de cisão não-trivial é $\mathbb{R} - \{0\} = (-\infty, 0) \cup (0, +\infty)$.

Um conjunto $X \subset \mathbb{R}^n$ chama-se *conexo* quando não admite outra cisão além da trivial. Assim, quando X é conexo, $X = A \cup B$, com A, B disjuntos e abertos em X, implica $A = \varnothing$ ou $B = \varnothing$.

O conjunto vazio e um ponto $\{x\}$ são exemplos óbvios de conjuntos conexos. Resulta do teorema de estrutura dos abertos da reta (vide Vol. 1, Cap. V, §1, Teorema 2) que todo intervalo aberto da reta é conexo. Em particular, $\mathbb{R}$ é conexo.

Quando existir uma cisão não-trivial $X = A \cup B$, diremos que X é *desconexo.*

Por exemplo, $\mathbb{R} - \{0\}$ é desconexo. Também é desconexo todo subconjunto discreto $X \subset \mathbb{R}^n$, com mais de um ponto. Com efeito, por definição, cada ponto $x \in X$ é isolado, isto é, dado $x \in X$, existe $\varepsilon > 0$ tal que $B(x;\varepsilon) \cap X = \{x\}$. Assim, cada ponto $x \in X$ é um conjunto aberto em X e, por conseguinte, todo subconjunto $A \subset X$ é aberto em X pois é reunião dos seus pontos. Obtém-se portanto uma cisão não trivial

$X = A \cup B$ considerando qualquer subconjunto $A \subset X$ com $\varnothing \neq A \neq X$ e pondo $B = X - A$.

O conjunto $\mathbb{Q}$ dos números racionais não é discreto mas um subconjunto conexo $X \subset \mathbb{Q}$ não pode ter mais de um ponto. Com efeito, se existirem $a, b \in X$ com $a < b$, escolheremos um número irracional ξ no intervalo aberto (a, b), poremos $A = (-\infty, \xi) \cap X$, $B = (\xi, +\infty) \cap X$ e obteremos a cisão $X = A \cup B$, que não é trivial porque $a \in A$ e $b \in B$.

Dada uma cisão $X = A \cup B$, temos $A = X - B$ e $B = X - A$, isto é, cada um dos conjuntos A, B é o complementar do outro em relação a X. Logo, cada um deles é aberto e fechado em X. Poderíamos portanto definir uma cisão de X como uma decomposição $X = A \cup B$, onde A e B são disjuntos e fechados em X. Poderíamos também definir um conjunto conexo X como aquele cujos únicos subconjuntos simultaneamente abertos e fechados em X são $\varnothing$ e o próprio X. Com efeito, dado $A \subset X$ aberto-fechado, $X = A \cup (X - A)$ é uma cisão.

Teorema 30. *A imagem de um conjunto conexo por uma aplicação contínua é um conjunto conexo.*

Demonstração: Sejam $X \subset \mathbb{R}^m$ conexo e $f\colon X \to \mathbb{R}^n$ contínua. Para toda cisão $f(X) = A \cup B$, o Teorema 16 (juntamente com a Observação que a ele se segue) garante que $X = f^{-1}(A) \cup f^{-1}(B)$ é uma cisão do conjunto conexo X. Logo $f^{-1}(A) = \varnothing$ (donde $A = \varnothing$) ou $f^{-1}(B) = \varnothing$ (donde $B = \varnothing$).

Corolário. *Todo conjunto homeomorfo a um conjunto conexo é também conexo.*

Teorema 31. *Um subconjunto $X \subset \mathbb{R}$ é conexo se, e somente se, é um intervalo.*

Demonstração: Já vimos que os intervalos abertos da reta são conexos. Os outros também são porque, dado qualquer intervalo $I \subset \mathbb{R}$, existe sempre uma função contínua $f\colon \mathbb{R} \to \mathbb{R}$ com $f(\mathbb{R}) = I$. Por exemplo, se $I = [-1, +1]$, tomamos $f(x) = \operatorname{sen} x$. Para $I = (0, 1]$, $f(x) = (1+x^2)^{-1}$. Se for $I = [0, +\infty)$, faremos $f(x) = x^2$. Todo intervalo não-aberto é homeomorfo a um desses três, logo é conexo. Reciprocamente, seja $X \subset R$ um subconjunto conexo. Se X não fosse um intervalo, existiriam $a < c < b$ com $a, b \in X$ e $c \notin X$. Então, pondo $A = (-\infty, c) \cap X$ e $B = (c, +\infty) \cap X$, obteríamos uma cisão $X = A \cup B$, a qual não seria trivial pois $a \in A$ e $b \in B$.

Corolário. *A imagem de um conjunto conexo $X \subset \mathbb{R}^n$ por uma função real contínua $f\colon X \to \mathbb{R}$ é um intervalo.*

Uma reformulação do Corolário acima é o

Teorema do Valor Intermediário. *Seja $f\colon X \to \mathbb{R}$ uma função contínua, definida num conjunto conexo $X \subset \mathbb{R}^n$. Se existem $a, b \in X$ e $d \in \mathbb{R}$ tais que $f(a) < d < f(b)$ então existe $c \in X$ tal que $f(c) = d$.*

Resulta do Teorema 30 que o círculo $S^1 = \{(x, y) \in \mathbb{R}^2; x^2 + y^2 = 1\}$ é conexo, pois é imagem da reta pela aplicação contínua $f\colon \mathbb{R} \to \mathbb{R}^2$, $f(t) = (\cos t, \operatorname{sen} t)$.

Como aplicação mostraremos que, dada $f\colon S^1 \to \mathbb{R}$ contínua, existe $u \in S^1$ tal que $f(u) = f(-u)$. Com efeito, a função $\varphi\colon S^1 \to \mathbb{R}$, definida por $\varphi(z) = f(z) - f(-z)$ é contínua no conexo S^1. Como $\varphi(-z) = -\varphi(z)$, resulta do Teorema do Valor Intermediário que deve existir $u \in S^1$ tal que $\varphi(u) = 0$, isto é, $f(u) = f(-u)$. Em particular, nenhuma função contínua $f\colon S^1 \to \mathbb{R}$ pode ser injetiva, portanto S^1 não é homeomorfa a um subconjunto da reta.

Daremos agora outra demonstração de que qualquer intervalo $I \subset \mathbb{R}$ é conexo, sem utilizar o teorema da estrutura dos abertos da reta.

Com efeito, suponhamos, por absurdo, que exista uma cisão não-trivial $I = A \cup B$ do intervalo I. Tomemos $a \in A$, $b \in B$, com $a < b$. Então $[a, b] \subset I$. Como A e B são fechados em I, os conjuntos $K = [a, b] \cap A$ e $L = [a, b] \cap B$ são fechados em $[a, b]$, logo são compactos. Além disso, $[a, b] = K \cup L$ é uma cisão. Pelo que vimos no parágrafo anterior, existem $x_0 \in K$ e $y_0 \in L$ tais que $|x_0 - y_0| = d(K, L)$. Isto significa que $|x_0 - y_0| \leq |x - y|$ sejam quais forem $x \in K$, $y \in L$. Note que $x_0 \neq y_0$ pois K e L são disjuntos. Tomemos o ponto médio c do intervalo cujos extremos são x_0 e y_0. Então $c \in [a, b]$. Em particular, $c \in K$ ou $c \in L$. Mas $|x_0 - c| < |x_0 - y_0|$, logo $c \notin L$. Também $|c - y_0| < |x_0 - y_0|$, donde $c \notin K$. Uma contradição.

Teorema 32 (Teorema da Alfândega). *Seja $X \subset \mathbb{R}^n$ um conjunto arbitrário. Se um conjunto conexo $C \subset \mathbb{R}^n$ contém um ponto $a \in X$ e um ponto $b \notin X$ então C contém algum ponto da fronteira de X.*

Demonstração: Se $a \in \partial X$ ou $b \in \partial X$, nada há para provar. Caso contrário, a função contínua $f\colon C \to \mathbb{R}$, definida por $f(x) = d(x, X) - d(x, R^n - X)$ é tal que $f(a) < 0$ e $f(b) > 0$. Pelo Teorema do Valor Intermediário, existe $x \in C$ tal que $f(x) = 0$. Então $d(x, X) = d(x, \mathbb{R}^n -$

X). Como pelo menos um desses números é zero, segue-se que ambos o são, ou seja, $x \in \partial X$.

Para uso nos dois teoremas seguintes, registremos a seguinte observação:

(14.1) *Sejam $X \subset Y \subset \mathbb{R}^n$. Se $A \subset Y$ é aberto em Y então $A \cap X$ é aberto em X.*

Com efeito, $X \subset Y$ significa $X = Y \cap X$ e A aberto em Y significa $A = B \cap Y$ com B aberto em $\mathbb{R}^n$. Logo $A \cap X = B \cap Y \cap X = B \cap X$ e assim $A \cap X$ é a interseção de X com o aberto B de $\mathbb{R}^n$, logo $A \cap X$ é aberto em X.

Teorema 33. *A reunião de uma família de conjuntos conexos com um ponto em comum é um conjunto conexo.*

Demonstração: Seja $X = \bigcup_{\lambda \in L} X_\lambda$ onde todos os conjuntos X_λ são conexos e contêm o ponto $a \in \mathbb{R}^n$. Para provar que X é conexo, suponhamos que $X = A \cup B$ seja uma cisão, digamos com $a \in A$. Para cada $\lambda \in L$, temos

$$X_\lambda = X \cap X_\lambda = (A \cup B) \cap X_\lambda = (A \cap X_\lambda) \cup (B \cap X_\lambda),$$

o que nos dá uma cisão de X_λ. Como X_λ é conexo e $a \in A \cap X_\lambda$, deve ser $B \cap X_\lambda = \varnothing$ para todo $\lambda \in L$. Logo

$$B = B \cap X = B \cap (\cup X_\lambda) = \bigcup_\lambda (B \cap X_\lambda) = \varnothing.$$

Assim, toda cisão de X é trivial e X é conexo.

Corolário 1. *A fim de que $X \subset \mathbb{R}^n$ seja conexo, é necessário e suficiente que, para quaisquer $a, b \in X$, exista um conjunto conexo C_{ab} com $a, b \in C_{ab} \subset X$.*

Fixando $a \in X$, teremos então $X = \bigcup_{x \in X} C_{ax}$, onde os conjuntos C_{ax} são conexos e têm o ponto a em comum. Pelo Teorema 33, X é conexo.

Corolário 2. *Dados $X \subset \mathbb{R}^m$ e $Y \subset \mathbb{R}^n$ não-vazios, o produto cartesiano $X \times Y \subset \mathbb{R}^{m+n}$ é conexo se, e somente se, X e Y são conexos.*

Com efeito, sejam X, Y conexos. Dados quaisquer $a = (a_1, a_2)$ e $b = (b_1, b_2)$ em $X \times Y$, seja $C_{ab} = (X \times a_2) \cup (b_1 \times Y)$. Temos $a \in C_{ab}$ e $b \in C_{ab}$. Além disso, $X \times a_2$ (homeomorfo a X) e $b_1 \times Y$ (homeomorfo

a Y) são conexos, com o ponto (b_1, a_2) em comum. Logo C_{ab} é conexo. Segue-se do Corolário 1 que $X \times Y$ é conexo. Reciprocamente, se o produto cartesiano $X \times Y$ é conexo então X e Y são conexos por serem imagens das projeções $\pi_1 \colon X \times Y \to X$ e $\pi_2 \colon X \times Y \to Y$.

Evidentemente, o Corolário 2 acima se estende para um produto cartesiano $X_1 \times X_2 \times \cdots \times X_n$ de um número finito qualquer de fatores. Em particular, o espaço euclidiano $\mathbb{R}^n = \mathbb{R} \times \cdots \times \mathbb{R}$ é conexo. Conseqüentemente, $\varnothing$ e $\mathbb{R}^n$ são os únicos subconjuntos de $\mathbb{R}^n$ que são ao mesmo tempo abertos e fechados.

Ainda em conseqüência do Teorema 33, toda bola (aberta ou fechada) do espaço euclidiano é conexa pois é reunião de seus raios, que são conexos por serem homeomorfos a intervalos da reta, e têm todos um ponto em comum, que é o centro da bola.

A interseção de conjuntos conexos pode não ser conexa. Por exemplo, os semi-círculos $X = \{(x, y) \in S^1; y \geq 0\}$ e $Y = \{(x, y) \in S^1; y \leq 0\}$ são conexos pois são ambos homeomorfos, através da projeção $(x, y) \mapsto x$, ao intervalo $[-1, 1]$. Mas a interseção $X \cap Y = \{e_1, -e_1\}$ é desconexa. Vale porém o

Teorema 34. *A interseção* $K = \bigcap_{i=1}^{\infty} K_i$ *de uma seqüência decrescente* $K_1 \supset K_2 \supset \cdots \supset K_i \supset \ldots$ *de conjuntos compactos conexos em* $\mathbb{R}^n$ *é um conjunto (compacto e) conexo.*

Demonstração: Seja $K = A \cup B$ uma cisão. Sendo fechados em K, os conjuntos A e B são fechados (e disjuntos) em $\mathbb{R}^n$. Existem portanto abertos disjuntos V, W em $\mathbb{R}^n$ tais que $A \subset V$ e $B \subset W$ (vide (13.3)). Assim, o aberto $U = V \cup W$ contém a interseção $K = \cap K_i$, logo contém algum dos conjuntos K_i (vide (12.8)). Então $K_i = (K_i \cap V) \cup (K_i \cap W)$ é uma cisão. Como K_i é conexo, deve-se ter $K_i \cap V = \varnothing$, donde $A = K \cap A \subset K_i \cap V = \varnothing$, ou então $K_i \cap W = \varnothing$, donde $B = \varnothing$. Portanto qualquer cisão de K é trivial, logo K é conexo.

Observe-se que se $F = \cap F_i$ for a interseção de uma seqüência decrescente $F_1 \supset F_2 \supset \cdots \supset F_i \ldots$ de subconjuntos fechados conexos (mas não compactos) então F pode ser desconexo. Basta tomar cada $F_i \subset \mathbb{R}^2$ como a reunião das duas retas horizontais $y = 0$, $y = 1$ mais os pontos (x, y) com $x \geq i$ e $0 \leq y \leq 1$.

Teorema 35. *Sejam* $X \subset Y \subset \overline{X}$ *em* $\mathbb{R}^n$. *Se* X *é conexo, então* Y *também é conexo.*

Demonstração: Como todo ponto de Y é aderente a X, para todo conjunto não-vazio A, aberto em Y, tem-se $A \cap X \neq \varnothing$. Feita esta observação, seja $Y = A \cup B$ uma cisão de Y. Então

$$X = Y \cap X = (A \cup B) \cap X = (A \cap X) \cup (B \cap X),$$

o que nos dá uma cisão de X. Como X é conexo, ou temos $A \cap X = \varnothing$ e portanto $A = \varnothing$ em virtude da observação inicial, ou $B \cap X = \varnothing$, donde $B = \varnothing$, pelo mesmo motivo. Assim, toda cisão de Y é trivial e Y é, portanto, conexo.

Corolário. *O fecho de um conjunto conexo é conexo.*

O Teorema 35 pode ser usado para mostrar que a esfera unitária $S^n = \{x \in \mathbb{R}^{n+1}; \langle x, x\rangle = 1\}$ é conexa. Com efeito, o conjunto $X = S^n - \{e_{n+1}\}$ é conexo, por ser homeomorfo a $\mathbb{R}^n$ através da projeção estereográfica (vide §8). Como e_{n+1} não é um ponto isolado em S^n, temos $\overline{X} = S^n$, logo S^n é conexa.

Outro exemplo é fornecido pelo conjunto $X \subset \mathbb{R}^2$, gráfico da função contínua $f\colon (0,1] \to \mathbb{R}$, $f(x) = \text{sen}(1/x)$. Por ser homeomorfo ao intervalo $(0,1]$, X é conexo. Seu fecho é $\overline{X} = X \cup I$, onde $I = \{(0,t) \in \mathbb{R}^2; -1 \leq t \leq 1\}$. Para todo subconjunto $T \subset I$, o Teorema 35 assegura que $X \cup T$ é conexo. Em particular, é conexo o conjunto $X_0 = X \cup \{p\}$, onde $p = (0,0)$ é a origem de $\mathbb{R}^2$. Este exemplo destoa da intuição, que nos sugere um conjunto conexo como aquele formado por um "só pedaço". Por causa disto, consideraremos a seguir uma noção mais estrita de conexidade: a conexidade por caminhos.

Um *caminho* num conjunto $X \subset \mathbb{R}^n$ é uma aplicação contínua $f\colon I \to X$, definida num intervalo I.

Por exemplo, dados $x, y \in \mathbb{R}^n$, o caminho $f\colon [0,1] \to \mathbb{R}^n$, definido por $f(t) = (1-t)x + ty$ chama-se o *caminho retilíneo* que liga x a y. Às vezes nos referiremos a ele como o caminho $[x, y]$.

Diremos que os pontos $a, b \in X$ *podem ser ligados por um caminho em* X quando existe um caminho $f\colon I \to X$ tal que $a, b \in f(I)$.

Por exemplo, se $X \subset \mathbb{R}^n$ é convexo, dois pontos quaisquer $a, b \in X$ podem ser ligados por um caminho em X, a saber, o caminho retilíneo $[a, b]$.

Se $a, b \in X$ podem ser ligados por um caminho $f\colon I \to X$, então existe um caminho $\varphi\colon [0,1] \to X$ tal que $\varphi(0) = a$ e $\varphi(1) = b$. Basta pôr $\varphi(t) = f((1-t)\alpha + t\beta)$, onde $a = f(\alpha)$ e $b = f(\beta)$.

Se $f, g\colon [0,1] \to X$ são caminhos em X, com $f(1) = g(0)$, então definimos o *caminho justaposto* $h = f \vee g\colon [0,1] \to X$ pondo $h(t) = f(2t)$ se $0 \leq t \leq 1/2$ e $h(t) = g(2t-1)$ se $1/2 \leq t \leq 1$. Note que estas duas expressões definem o mesmo valor de $h(1/2)$. Como $h \mid [0,1/2]$ e $h \mid [1/2,1]$ são contínuas, segue-se que h é contínua. Intuitivamente, o caminho h percorre a trajetória de f (com velocidade dobrada) até $t = 1/2$ e depois, para $t \geq 1/2$, descreve (ainda com velocidae dobrada) o percurso de g.

Sejam a, b, c pontos do conjunto $X \subset \mathbb{R}^n$. Se a, b podem ser ligados por um caminho em X e b, c também podem ser ligados por um caminho em X, então existe um caminho em X ligando a a c. Basta tomar caminhos $f, g\colon [0,1] \to X$ com $f(0) = a,\ \ f(1) = b,\ \ g(0) = b,\ \ g(1) = c$ e pôr $h = f \vee g$. Então $h(0) = a,\ \ h(1) = c$.

Um conjunto $X \subset \mathbb{R}^n$ diz-se *conexo por caminhos* quando dois pontos quaisquer $a, b \in X$ podem ser ligados por um caminho em X.

Todo conjunto convexo $X \subset \mathbb{R}^n$ é conexo por caminhos. Em particular, toda bola (aberta ou fechada) no espaço euclidiano é conexa por caminhos.

A esfera $S^n = \{x \in \mathbb{R}^{n+1}; \langle x, x\rangle = 1\}$ é conexa por caminhos. Com efeito, dados $a, b \in S^n$, se a e b não são antípodas, isto é, se $b \neq -a$, então $f\colon [0,1] \to S^n$, definida por

$$f(t) = \frac{(1-t)a + tb}{|(1-t)a + tb|}$$

é contínua (pois seu denominador nunca se anula), com $f(0) = a$, $f(1) = b$. Se, porém, $b = -a$, tomamos um ponto $c \in S^n - \{a, b\}$, ligamos a com c e c com b pelo processo acima. O caminho justaposto ligará o ponto a ao seu antípoda b.

Todo conjunto conexo por caminhos é conexo, em virtude do Corolário 1 do Teorema 33, pois se $f\colon I \to X$ é um caminho em X ligando os pontos a e b então $f(I) = C_{ab}$ é um subconjunto conexo de X contendo a e b.

A recíproca é falsa. O conjunto $X_0 \subset \mathbb{R}^2$, acima descrito como a reunião do gráfico da função $f(x) = \operatorname{sen}(1/x)$, $\ 0 < x \leq 1$, com a origem $p = (0,0)$, é conexo mas não é conexo por caminhos. (Para a demonstração, ver o livro “Espaços Métricos”, do autor, página 103.)

Há, porém um caso particular importante, no qual a conexidade implica em conexidade por caminhos: quando o conjunto $X \subset \mathbb{R}^n$ é aberto.

Diremos que $f\colon [0,1] \to X$ é um *caminho poligonal* em X quando f é a justaposição de um número finito de caminhos retilíneos.

Teorema 36. *Um aberto $A \subset \mathbb{R}^n$ é conexo se, e somente se, é conexo por caminhos.*

Demonstração: Seja $A \subset \mathbb{R}^n$ aberto e conexo. Fixemos um ponto $a \in A$ e consideremos o conjunto U, formado pelos pontos $x \in A$ que podem ser ligados ao ponto a por um caminho poligonal contido em A. Afirmamos que U é aberto. Com efeito, seja $x \in U$. Sendo A aberto, existe $B = B(x;r)$, com $x \in B \subset A$. Como a bola B é convexa, todo ponto $y \in B$ pode ser ligado a x por um segmento de reta contido em B, logo y se liga a a por um caminho poligonal contido em A. Logo $B \subset U$ e $U \subset A$ é aberto. Também $V = A - U$ é aberto, pois se $v \in V$ então v não pode ser ligado a a por um caminho poligonal contido em A. Tomando uma bola aberta B_1, com $v \in B_1 \subset A$, todo $z \in B_1$ se liga a v por um segmento de reta contido em B_1. Se z pudesse ser ligado a a por um caminho poligonal contido em A, justapondo-se $[v,z]$ a esse caminho, veríamos que $v \in U$, um absurdo. Temos então $A = U \cup V$, uma cisão. Como A é conexo e $a \in U$, temos $V = \varnothing$, donde $A = U$, o que prova o teorema.

Corolário (da demonstração). *Se $A \subset \mathbb{R}^n$ é aberto e conexo, dois pontos quaisquer de A podem ser ligados por um caminho poligonal contido em A.*

O problema central da topologia dos espaços euclidianos é determinar se dois conjuntos $X \subset \mathbb{R}^m$, $Y \subset \mathbb{R}^n$, dados são ou não homeomorfos. Não existe uma resposta geral para este problema. Para afirmar que X e Y são homeomorfos é necessário definir um homeomorfismo entre eles. Para garantir que X e Y não são homeomorfos, procura-se lançar mão de invariantes topológicos como a compacidade, a conexidade (e outros mais elaborados, que não estudaremos aqui).

Vejamos um exemplo. Sejam $X=\{(x,y)\in\mathbb{R}^2; x^2+y^2=1\}$ um círculo, $Y=\{(x,y)\in\mathbb{R}^2; 4x^2+9y^2=1\}$ uma elipse, $Z=\{(x,y)\in\mathbb{R}^2; x^2-y^2=1\}$ uma hipérbole e $T=\{(x,y)\in\mathbb{R}^2; y=x^2\}$ uma parábola. X e Y são homeomorfos e $h\colon X \to Y$, $h(x,y)=(x/2,y/3)$, é um homeomorfismo entre eles. Mas, como X, Y são compactos e Z, T não são, segue-se que não pode existir um homeomorfismo entre um desses dois últimos conjuntos e um dos dois primeiros. Por outro lado, Z e T não são homeomorfos porque T é conexo e Z não é.

Sejam agora $X = [a, b]$ um intervalo fechado da reta e $Y = B[c; r] \subset \mathbb{R}^2$ o disco fechado de centro $c \in \mathbb{R}^2$ e raio r. Ambos X, Y são compactos e conexos; no entanto a intuição nos diz que X não deve ser homeomorfo a Y. De fato, não é e a prova se baseia na seguinte observação: seja qual for o ponto $y \in Y$, o conjunto $Y - \{y\}$ ainda é conexo enquanto que, para $a < x < b$, o conjunto $X - \{x\}$ é desconexo. Se existisse um homeomorfismo $h \colon X \to Y$, escolheríamos um ponto $x \in (a, b)$, escreveríamos $y = h(x)$ e teríamos, por restrição, um homeomorfismo $k \colon X - \{x\} \to Y - \{y\}$, $k = h \mid (X - \{x\})$, entre um conjunto desconexo e um conjunto conexo.

Se tentarmos usar esse raciocínio para provar que um disco $X = B[a; r] \subset \mathbb{R}^2$ não é homeomorfo a uma bola fechada $Y = B[b; s] \subset \mathbb{R}^3$ não chegaremos a nada porque X e Y permanecem conexos depois da retirada de qualquer um dos seus pontos. É verdade que uma bola no espaço euclidiano $\mathbb{R}^m$ só é homeomorfa a uma bola em $\mathbb{R}^n$ quando $m = n$. Mas a demonstração deste fato requer o uso dos invariantes topológicos "mais elaborados"a que nos referimos acima, os quais são estudados em Topologia Algébrica ou Topologia Diferencial.

Voltando a uma situação mais elementar, podemos constatar que não existe um homeomorfismo entre o conjunto $X = \{(x, y) \in \mathbb{R}^2; x^2 = y^2\}$ (um par de retas que se cortam na origem) e a parábola $Y = \{(x, y) \in \mathbb{R}^2; y = x^2\}$. Com efeito, a retirada de qualquer ponto de Y deixa um conjunto com "dois pedaços"enquanto que se retirarmos a origem de X ficaremos com "quatro pedaços". Este tipo de argumento lança mão dos pedaços conexos que compõem um conjunto desconexo. Tornaremos precisa esta idéia mediante a noção de componente conexa, que definiremos agora. Mostraremos a seguir que todo conjunto $X \subset \mathbb{R}^n$ se exprime como reunião disjunta de subconjuntos conexos máximos, chamados componentes conexas de X.

Sejam $x \in X \subset \mathbb{R}^n$. A *componente conexa* do ponto x no conjunto X é a reunião C_x de todos os subconjuntos conexos de X que contêm o ponto x.

Por exemplo, se $X = \mathbb{Q} \subset \mathbb{R}$ então a componente conexa de qualquer ponto $x \in X$ é $\{x\}$. Por outro lado, se $X \subset \mathbb{R}^n$ é conexo então, para todo $x \in X$ temos $C_x = X$. Se $X = \mathbb{R} - \{0\}$ então a componente conexa de 1 em X é $(0, +\infty)$ enquanto que a componente conexa de -1 é $(-\infty, 0)$.

Dados $x \in X \subset \mathbb{R}^n$, a componente conexa C_x é um conjunto conexo, pelo Teorema 33. Na realidade, C_x é o maior subconjunto conexo de X contendo o ponto x. Com efeito, se $C \subset X$ é conexo e contém x, então C é um dos conjuntos cuja reunião é C_x, logo $C \subset C_x$. Mais ainda, se $C \subset X$ é conexo e tem algum ponto em comum com C_x então $C \subset C_x$, pois $C \cup C_x$ é conexo contendo x logo $C \cup C_x \subset C_x$ e daí $C \subset C_x$. Em particular, nenhum subconjunto conexo de X pode conter C_x propriamente.

Sejam x, y dois pontos de X. Suas componentes conexas C_x e C_y ou coincidem ou são disjuntas pois se $z \in C_x \cap C_y$ então $C_x \subset C_y$ e $C_y \subset C_x$. Assim a relação "x e y pertencem à mesma componente conexa em X"é uma equivalência no conjunto X. As classes de equivalência são as componentes conexas dos pontos de X.

Toda componente conexa C_x é um conjunto fechado em X. Com efeito, sendo $C_x \subset \overline{C}_x \cap X \subset \overline{C}_x$, o Teorema 35 nos assegura que $\overline{C}_x \cap X$ é um subconjunto conexo de X, contendo C_x. Logo $\overline{C}_x \cap X = C_x$, o que mostra que C_x é fechado em X.

Finalmente, seja $h\colon X \to Y$ um homeomorfismo. Se C_x é a componente conexa do ponto x em X, então $D_y = h(C_x)$ é a componente conexa do ponto $y = h(x)$ em Y. Com efeito, pelo Teorema 30, D_y é conexo. Além disso se $D \subset Y$ é qualquer conexo contendo y então, como $h^{-1}\colon Y \to X$ é contínua, $h^{-1}(D)$ é um subconjunto conexo de X, contendo x. Logo $h^{-1}(D) \subset C_x$ e daí $D \subset h(C_x) = D_y$. Portanto D_y é o maior subconjunto conexo de Y contendo y, ou seja, é a componente conexa de y em Y. O homeomorfismo $h\colon X \to Y$ estabelece portanto uma bijeção entre as componentes conexas de X e as componentes conexas de Y.

15 A norma de uma transformação linear

Fixemos arbitrariamente uma norma em $\mathbb{R}^m$ e outra em $\mathbb{R}^n$. Sabemos que toda transformação linear $A\colon \mathbb{R}^m \to \mathbb{R}^n$ é contínua. Mais precisamente, existe um $c > 0$ tal que $|A \cdot x| \leq c|x|$ para todo $x \in \mathbb{R}^m$. Nesta desigualdade, a norma à direita é a de $\mathbb{R}^m$, enquanto a da esquerda é a de $\mathbb{R}^n$. Elas podem ser diferentes, mesmo quando $m = n$. Tudo depende da nossa escolha.

Se nos restringirmos a considerar vetores unitários em $\mathbb{R}^m$, resulta que $x \in \mathbb{R}^m$, $|x| = 1 \Rightarrow |A \cdot x| \leq c$, isto é, A transforma a esfera unitária de $\mathbb{R}^m$ num subconjunto limitado de $\mathbb{R}^n$. Isto também poderia ser visto observando que a esfera unitária de $\mathbb{R}^m$ é compacta, logo é transformada pela aplicação contínua A num subconjunto compacto – em particular, limitado – de $\mathbb{R}^n$.

Estas considerações nos permitem introduzir uma norma no espaço vetorial $\mathcal{L}(\mathbb{R}^m; \mathbb{R}^n) = \mathbb{R}^{nm}$, pondo, para cada transformação linear $A\colon \mathbb{R}^m \to \mathbb{R}^n$,

$$|A| = \sup\{|A \cdot x|; x \in \mathbb{R}^m, |x| = 1\}.$$

Deixaremos ao leitor o trabalho de verificar que a função $A \mapsto |A|$ cumpre as condições exigidas para uma norma, além de gozar das duas propriedades seguintes:

$$|A \cdot x| \leq |A|\,|x| \text{ para todo } x \in \mathbb{R}^m$$

$$|A \cdot B| \leq |A|\,|B| \text{ se } A \in \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n) \text{ e } B \in \mathcal{L}(\mathbb{R}^k; \mathbb{R}^m).$$

Observação: Na segunda desigualdade acima, a norma em $\mathbb{R}^m$ deve ser tomada a mesma, tanto no espaço $\mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$ como em $\mathcal{L}(\mathbb{R}^k; \mathbb{R}^m)$.

Como duas normas quaisquer em $\mathbb{R}^{nm}$ são equivalentes, dada uma seqüência de transformações lineares $A_k\colon \mathbb{R}^m \to \mathbb{R}^n$ e uma transformação $A \in \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$, temos $\lim\limits_{k\to\infty} |A_k - A| = 0$ se, e somente se, cada elemento da matriz de A_k converge para o elemento correspondente da matriz de A.

Para certas escolhas de normas em $\mathbb{R}^m$ e $\mathbb{R}^n$, pode-se determinar a norma $|A|$ em função dos elementos a_{ij} da matriz de A. Às vezes, o melhor que se tem é uma desigualdade que dá um limite superior para $|A|$. Na tabela seguinte, examinamos o que ocorre com as três normas usuais do espaço euclidiano. Nas fórmulas, o índice i varia de 1 a n, enquanto $1 \leq j \leq m$.

Norma em $\mathbb{R}^m$	Norma em $\mathbb{R}^n$	Norma de $A\colon \mathbb{R}^m \to \mathbb{R}^n$
do máximo	do máximo	$\lvert A\rvert = \max_i(\sum_j \lvert a_{ij}\rvert)$ maior "norma da soma" entre as linhas
da soma	da soma	$\lvert A\rvert = \max_j(\sum_i \lvert a_{ij}\rvert)$ maior "norma da soma" entre as colunas
euclidiana	euclidiana	$\lvert A\rvert = \sqrt{\lambda}, \lambda =$ maior autovalor de A^*A
da soma	do máximo	$\lvert A\rvert = \max_{i,j} \lvert a_{ij}\rvert$
		$\lvert A\rvert \le \sum_{i,j} \lvert a_{ij}\rvert$
do máximo	da soma	(vide valor exato de $\lvert A\rvert$ na última linha)
euclidiana	do máximo	$\lvert A\rvert = \max_i \sqrt{\sum_j a_{ij}^2}$
euclidiana	da soma	$\lvert A\rvert \le \sum_i(\sqrt{\sum_j a_{ij}^2})$
do máximo	euclidiana	$\lvert A\rvert \le m\sqrt{\sum_i \max_j (a_{ij})^2}$
da soma	euclidiana	$\lvert A\rvert \le \sqrt{\sum_i \max_j (a_{ij})^2}$
euclidiana	$n = 1$ Neste caso as 3 normas coincidem	$f(x) = a_1x_1 + \cdots + a_mx_m \Rightarrow$ $\Rightarrow \lvert f\rvert = \sqrt{\sum_j a_j^2}$
euclidiana, A simétrica	euclidiana, $n = m$	$\lvert A\rvert = \max\{\lvert\lambda_1\rvert, \ldots, \lvert\lambda_n\rvert\}$ onde $\lambda_1, \ldots, \lambda_n$ são os autovalores de A
(repetição) do máximo	da soma	$\lvert A\rvert = \max_\varepsilon \sum_i \lvert \sum_j \varepsilon_j a_{ij}\rvert$, onde $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_m)$ é tal que $\varepsilon_j = \pm 1$

Exercícios

§1.

1.1. Mostre que as operações usuais de soma de aplicações e produto de uma aplicação por um número real fazem do conjunto $\mathcal{L}(\mathbb{R}^m;\mathbb{R}^n)$ um espaço vetorial. Analogamente para o conjunto $M(n \times m)$. Mostre que as bijeções estabelecidas no texto entre esses conjuntos e $\mathbb{R}^{nm}$ são isomorfismos entre espaços vetoriais. Exiba explicitamente bases para os espaços $\mathcal{L}(\mathbb{R}^m;\mathbb{R}^n)$ e $M(n \times m)$.

1.2. Seja $E = \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n; \mathbb{R}^p)$ o conjunto das aplicações bilineares $\varphi\colon \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^p$. Mostre que as operações usuais fazem de E um espaço vetorial de dimensão mnp.

1.3. Seja E o espaço vetorial das funções bilineares $\varphi\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$. Estabeleça um isomorfismo entre E e o espaço vetorial $M(n \times n)$ das matrizes reais $n \times n$. Defina função bilinear simétrica e mostre que tal isomorfismo leva funções bilineares simétricas em matrizes simétricas. Mostre que a matriz correspondente à função bilinear φ é invertível se, e somente se, φ é não-degenerada (isto é, $\varphi(x, y) = 0$ para todo $y \in \mathbb{R}^n \Rightarrow x = 0$).

1.4. Seja $E \subset \mathbb{R}^n$ um subespaço vetorial de dimensão m. Prove que existem $n - m$ funcionais lineares $f_1, f_2, \ldots, f_{n-m}\colon \mathbb{R}^n \to \mathbb{R}$ tais que $E = \{x \in \mathbb{R}^n; f_1(x) = f_2(x) = \cdots = f_{n-m}(x) = 0\}$. Conclua que existe uma aplicação linear sobrejetiva $A\colon \mathbb{R}^n \to \mathbb{R}^{n-m}$ tal que $E = A^{-1}(0)$.

§2.

2.1. Para todo funcional linear $f \in (\mathbb{R}^n)^*$ existe um único vetor $y \in \mathbb{R}^n$ tal que $f(x) = \langle y, x\rangle$ qualquer que seja $x \in \mathbb{R}^n$.

2.2. Um conjunto $\{u_1, \ldots, u_r\} \subset \mathbb{R}^n$ diz-se *ortonormal* quando $\langle u_j, u_j\rangle = 1$ e $\langle u_i, u_j\rangle = 0$ para $i \neq j$ quaisquer. Todo conjunto ortonormal é parte de uma base ortonormal. Se $\{u_1, \ldots, u_n\}$ é uma base ortonormal então $x = \sum_{i=1}^{n} \langle x, u_i\rangle u_i$ para todo $x \in \mathbb{R}^n$.

2.3. Considere em $\mathbb{R}^m$ e em $\mathbb{R}^n$ a norma euclidiana. Dada uma aplicação linear $A\colon \mathbb{R}^m \to \mathbb{R}^n$, existe uma única aplicação linear $A^*\colon \mathbb{R}^n \to \mathbb{R}^m$, chamada a *adjunta* de A, tal que $\langle Ax, y\rangle = \langle x, A^*y\rangle$ para quaisquer $x \in \mathbb{R}^m$, $y \in \mathbb{R}^n$. Dado $b \in \mathbb{R}^n$, a equação $Ax = b$ possui solução $x \in \mathbb{R}^m$ se, e somente se, b é ortogonal a todo elemento do núcleo de A^*. Conclua que a imagem de A^* e a imagem de A têm a mesma dimensão.

2.4. Uma aplicação linear $A\colon \mathbb{R}^n \to \mathbb{R}^n$ diz-se *simétrica* quando $A = A^*$. Prove qe o conjunto S das aplicações lineares simétricas constitui um subespaço vetorial de dimensão $n(n+1)/2$ em $\mathcal{L}(\mathbb{R}^n;\mathbb{R}^n)$. Quando $A^* = -A$, diz-se que A é *anti-simétrica.* Prove que o conjunto T das aplicações lineares anti-simétricas é um subespaço vetorial de dimensão $n(n-1)/2$ em $\mathcal{L}(\mathbb{R}^n;\mathbb{R}^n)$ e que toda aplicação linear A se escreve, de modo único, como soma de uma aplicação simétrica com uma anti-simétrica, isto é, $\mathcal{L}(\mathbb{R}^n;\mathbb{R}^n) = S \oplus T$.

2.5. Considere em $\mathbb{R}^m$ e $\mathbb{R}^n$ a norma euclidiana. As seguintes afirmações a respeito de uma aplicação linear $A\colon \mathbb{R}^m \to \mathbb{R}^n$ são equivalentes:

(i) $|Ax| = |x|$ para todo $x \in \mathbb{R}^m$;

(ii) $|Ax - Ay| = |x - y|$ para quaisquer $x, y \in \mathbb{R}^m$;

(iii) $\langle Ax, Ay \rangle = \langle x, y \rangle$ para quaisquer $x, y \in \mathbb{R}^m$;

(iv) Todo conjunto ortonormal em $\mathbb{R}^m$ é transformado por A num conjunto ortonormal em $\mathbb{R}^n$;

(v) $A^*A = I_m$ (aplicação identidade de $\mathbb{R}^m$);

(vi) As colunas da matriz de A formam um conjunto ortonormal em $\mathbb{R}^n$.

Quando $m = n$, tem-se também $AA^* = I_m$ e a aplicação linear A chama-se *ortogonal.*

2.6. Se A é ortogonal então $\det A = \pm 1$.

2.7. Dados os números reais a, b, c, a fim de que exista em $\mathbb{R}^2$ um produto interno tal que $\langle e_1, e_1 \rangle = a$, $\langle e_1, e_2 \rangle = \langle e_2, e_1 \rangle = b$ e $\langle e_2, e_2 \rangle = c$, é necessário e suficiente que $a > 0$ e $ac > b^2$.

2.8. Existe em $\mathbb{R}^3$ um produto interno tal que $\langle e_1, e_1 \rangle = 2$, $\langle e_2, e_2 \rangle = 3$, $\langle e_3, e_3 \rangle = 4$, $\langle e_1, e_2 \rangle = 0$ e $\langle e_2, e_3 \rangle = \langle e_1, e_3 \rangle = 1$.

2.9. Se $c \in [a, b]$ então $|b - a| = |b - c| + |c - a|$. Se a norma provém de um produto interno, vale a recíproca. Para uma norma arbitrária pode-se ter a igualdade acima com $c \notin [a, b]$.

2.10. Se a norma provém de um produto interno e $a \neq b$ em $\mathbb{R}^n$ são tais que $|a| \leq r$ e $|b| \leq r$ então $|(1 - t)a + tb| < r$ para todo $t \in (0, 1)$. (Em particular a esfera não contém segmentos de reta.)

2.11. Seja $C \subset \mathbb{R}^n$ um conjunto convexo. Fixado $p \in \mathbb{R}^n$, seja $\varphi\colon C \to \mathbb{R}$ a função definida por $\varphi(x) = |x - p| = \sqrt{\langle x - p, x - p \rangle}$. Existe no máximo um ponto $a \in C$ tal que $\varphi(a) = \inf\{\varphi(x); x \in C\}$.

2.12. Fixe números reais α, β com $\alpha < \beta$ e, para cada $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ ponha $|x| = \sup\limits_{\alpha \leq t \leq \beta} |x_1 + x_2 t + \cdots + x_n t^{n-1}|$. Prove que isto define uma norma em $\mathbb{R}^n$, a qual não provém de um produto interno.

2.13. Dado um subconjunto $X \subset \mathbb{R}^n$, seu *complemento ortogonal* é o conjunto $X^\perp = \{y \in \mathbb{R}^n; \langle x, y \rangle = 0$ para todo $x \in X\}$, $X^\perp$ é um subespaço vetorial de $\mathbb{R}^n$. Se $E \subset \mathbb{R}^n$ é um subespaço vetorial então $E^{\perp\perp} = E$.

2.14. Sejam $E \subset \mathbb{R}^n$ um subespaço vetorial, $(e_1, \ldots, e_k)$ uma base ortonormal de E e $a \in \mathbb{R}^n$ um vetor arbitrário. Pondo $a_0 = \langle a, e_1 \rangle e_1 + \cdots + \langle a, e_k \rangle e_k$, o vetor $a - a_0$ é perpendicular a todos os vetores de E, e $|a - a_0| \leq |a - y|$ para todo $y \in E$. A função $\varphi\colon E \to \mathbb{R}$, definida por $\varphi(y) = |a - y|$, atinge seu valor mínimo num único ponto de E, a saber, o ponto a_0. Daí resulta que a_0 depende apenas de a, mas não da base ortonormal escolhida em E. A aplicação $\pi\colon \mathbb{R}^n \to E$, dada por $\pi(a) = a_0$, é linear, seu núcleo é $E^\perp$ e todo vetor $z \in \mathbb{R}^n$ se escreve, de modo único, como $z = x + y$, com $x \in E$ e $y \in E^\perp$, logo $\mathbb{R}^n = E \oplus E^\perp$.

2.15. Prove que as seguintes afirmações a respeito de uma aplicação linear $A\colon \mathbb{R}^n \to \mathbb{R}^n$ são equivalentes:

(i) Existe $\alpha > 0$ tal que $\langle Ax, Ay\rangle = \alpha^2\langle x, y\rangle$, quaisquer que sejam $x, y \in \mathbb{R}^n$;

(ii) $|Ax| = \alpha \cdot |x|$ para todo $x \in \mathbb{R}^n$ (α constante);

(iii) Se $(e_1, \ldots, e_n)$ é uma base ortonormal, então $\langle Ae_i, Ae_j\rangle = 0$ para $i \neq j$ e $|Ae_i| = \alpha$ para quaisquer $i, j = 1, \ldots, n$.

Quando isto ocorre, A chama-se uma *semelhança*.

2.16. A fim de que $A\colon \mathbb{R}^n \to \mathbb{R}^n$ seja uma semelhança, é necessário e suficiente que exista $\alpha > 0$ tal que $(1/\alpha) \cdot A$ seja ortogonal.

§3.

3.1. O conjunto das matrizes reais 2×2 da forma $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ constitui um corpo, que é isomorfo ao corpo $\mathbb{C}$ dos números complexos.

3.2. Dados $z, w \in \mathbb{C} - \{0\}$, tem-se $\measuredangle(w, z) = \measuredangle\left(\frac{1}{z}, \frac{1}{w}\right)$.

3.3. Dados $a = re^{it} \in \mathbb{C} - \{0\}$ e $n \in \mathbb{N}$, a equação $z^n - a = 0$ possui exatamente n raízes, que são os números complexos

$$z_j = \sqrt[n]{r}\left(\cos\frac{t + 2\pi j}{n} + i \operatorname{sen}\frac{t + 2\pi j}{n}\right),$$

$j = 0, 1, \ldots, n - 1$.

3.4. As seguintes afirmações a respeito de uma aplicação linear $A\colon \mathbb{R}^2 \to \mathbb{R}^2$ são equivalentes:

(i) A é uma semelhança e $\det A > 0$;

(ii) Existe um número complexo $a \neq 0$ tal que $A \cdot z = a \cdot z$ (multiplicação complexa) para todo $z \in \mathbb{R}^2$.

§4.

4.1. Qualquer que seja a norma adotada em $\mathbb{R}^n$ $(n > 1)$, a esfera unitária $S = \{x \in \mathbb{R}^n; |x| = 1\}$ é um conjunto infinito.

4.2. Dados $x \in S[a; r]$ e $\varepsilon > 0$, prove que existem $y \in B(a; r)$ e $z \notin B[a; r]$ tais que $|y - x| < \varepsilon$ e $|z - x| < \varepsilon$.

4.3. Se $X \subset \mathbb{R}^m$ e $Y \subset \mathbb{R}^n$ são convexos então seu produto cartesiano $X \times Y \subset \mathbb{R}^{m+n}$ é convexo.

4.4. A interseção de uma família arbitrária de conjuntos convexos é um conjunto convexo.

4.5. Dados $X, Y \subset \mathbb{R}^n$, seja $X * Y$ a reunião de todos os segmentos de reta $[x, y]$, onde x varia em X e y em Y. Se X e Y são convexos então $X * Y$ é convexo.

4.6. Dados $X \subset \mathbb{R}^n$ e $\varepsilon > 0$, seja $B(X; \varepsilon)$ a reunião das bolas $B(x; \varepsilon)$ com $x \in X$. Se X é convexo então $B(X; \varepsilon)$ é convexo.

4.7. Dado $X \subset \mathbb{R}^n$, a *envoltória convexa* de X é a interseção $C(X)$ de todos os subconjuntos convexos de R^n que contêm X. Prove que $C(X)$ é o conjunto de todas as combinações lineares $\alpha_1 x_1 + \cdots + \alpha_k x_k$ tais que $x_1, \ldots, x_k \in X$, $\alpha_1 \geq 0, \ldots, \alpha_k \geq 0$ e $\alpha_1 + \cdots + \alpha_k = 1$.

§5.

5.1. Se existirem seqüências de pontos $x_k, y_k \in \mathbb{R}^n$ com $\lim x_k = a$, $\lim y_k = b$ e $|y_k - a| < r < |x_k - b|$ para todo $k \in \mathbb{N}$ então $|a - b| = r$.

5.2. As seguintes afirmações a respeito de uma seqüência (x_k) de pontos de $\mathbb{R}^n$ são equivalentes:

(a) $\lim\limits_{k\to\infty} |x_k| = +\infty$;

(b) (x_k) não possui subseqüência convergente;

(c) Para todo conjunto limitado $L \subset \mathbb{R}^n$, o conjunto dos índices k tais que $x_k \in L$ é finito.

5.3. Se $b \in B(a;r) \subset \mathbb{R}^n$ e $\lim x_k = b$ então existe $k_0 \in \mathbb{N}$ tal que $k > k_0 \Rightarrow x_k \in B(a;r)$.

5.4. Defina convergência e convergência absoluta (ou normal) de uma série Σx_k cujos termos $x_k = (x_{k1}, \ldots, x_{kn})$ pertencem a $\mathbb{R}^n$. Prove que a série Σx_k converge (resp. converge absolutamente) se, e somente se, para cada $i = 1, \ldots, n$, a série $\sum\limits_k x_{ki}$ converge (resp. converge absolutamente). Conclua que toda série absolutamente convergente em $\mathbb{R}^n$ é convergente.

5.5. Prove que $\lim x_k = a$ em $\mathbb{R}^n$ se, e somente se, $\lim \langle x_k, y\rangle = \langle a, y\rangle$ para todo $y \in \mathbb{R}^n$.

5.6. Dada uma seqüência de aplicações lineares $A_k\colon \mathbb{R}^m \to \mathbb{R}^n$ suponha que, para todo $x \in \mathbb{R}^m$, exista $Ax = \lim\limits_{k\to\infty} A_k x$. Prove que a aplicação $A\colon \mathbb{R}^m \to \mathbb{R}^n$, assim definida, é linear, que $\lim\limits_{k\to\infty} A_k = A$ relativamente a qualquer norma em $\mathcal{L}(\mathbb{R}^m;\mathbb{R}^n)$ e que a convergência $A_k x \to Ax$ é uniforme em toda parte limitada de $\mathbb{R}^m$.

5.7. Para toda aplicação $X \in \mathcal{L}(R^n;R^n)$, a série $\sum\limits_{k=0}^{\infty} X^k/k!$ é absolutamente convergente. Indicando sua soma por e^X, tem-se $e^X \cdot e^Y = e^{X+Y}$ desde que $XY = YX$. Conclua que, para toda $X \in \mathcal{L}(\mathbb{R}^n;\mathbb{R}^n)$, e^X é invertível, com $(e^X)^{-1} = e^{-X}$.

5.8. Toda matriz $n \times n$ é limite de uma seqüência de matrizes invertíveis $n \times n$.

§6.

6.1. As seguintes afirmações a respeito de um ponto $a \in \mathbb{R}^n$ e um conjunto $X \subset R^n$ são equivalentes:

(i) $a = \lim x_k$ com $x_k \in X$ e $x_k \neq a$ para todo $k \in \mathbb{N}$;

(ii) $a = \lim y_k$ com $y_k \in X$ para todo $k \in \mathbb{N}$ e $y_k \neq y_\ell$ se $k \neq \ell$.

6.2. Se nenhum ponto do conjunto X é ponto de acumulação então se pode escolher, para cada ponto $x \in X$, uma bola aberta B_x , de centro x, de tal maneira que, para $x \neq y$ em X se tenha $B_x \cap B_y = \varnothing$.

6.3. Todo conjunto discreto é enumerável. Noutras palavras: todo conjunto não-enumerável contém um ponto de acumulação.

§7.

7.1. Seja $f\colon X \to \mathbb{R}^n$ contínua. Dada uma seqüência de pontos $x_k \in X$ com $\lim x_k = a \in X$ e $|f(x_k)| \leq c$ para todo $k \in \mathbb{N}$ então $|f(a)| \leq c$.

7.2. Sejam $f, g\colon X \to \mathbb{R}^n$ contínuas no ponto $a \in X$. Se $f(a) \neq g(a)$ então existe uma bola B de centro a tal que $x, y \in B \Rightarrow f(x) \neq g(y)$.

7.3. Seja $f\colon X \to \mathbb{R}^n$ contínua no ponto $a \in X$. Se $f(a)$ não pertence à bola fechada $B[b; r] \subset \mathbb{R}^n$ então existe $\delta > 0$ tal que $x \in X$, $|x - a| < \delta \Rightarrow f(x) \notin B[b; r]$.

7.4. Seja $f\colon X \to \mathbb{R}$ contínua no ponto $a \in X$. Se $f(a) > 0$ então existe $\delta > 0$ tal que $x \in X$, $|x - a| < \delta \Rightarrow f(x) > 0$.

7.5. Seja $f\colon \mathbb{R}^m \to \mathbb{R}^n$ contínua. Se $X \subset \mathbb{R}^m$ é limitado então $f(X) \subset \mathbb{R}^n$ é limitado.

7.6. Uma aplicação $f\colon X \to \mathbb{R}^n$ chama-se *localmente Lipschitziana* quando, para todo $x \in X$, existe uma bola aberta B, de centro x, tal que $f|(B \cap X)$ é Lipschitziana. Prove que a função $f\colon (0, 1] \to \mathbb{R}$, dada por $f(x) = \sqrt{x}$, é localmente Lipschitziana. Mostre que a função $g\colon [0, 1] \to \mathbb{R}$, $g(x) = \sqrt{x}$, embora contínua, não é localmente Lipschitziana.

7.7. Se a aplicação linear $A\colon \mathbb{R}^m \to \mathbb{R}^n$ é injetiva então existe $c > 0$ tal que $|Ax| \geq c|x|$ para todo $x \in \mathbb{R}^m$.

7.8. Seja B a bola aberta de centro na origem e raio 1 em $\mathbb{R}^n$. A aplicação contínua $f\colon B \to \mathbb{R}^n$, definida por $f(x) = x/(1 - |x|)$, não é uniformemente contínua.

7.9. Se $f\colon \mathbb{R}^m \to \mathbb{R}^n$ é contínua então, para cada parte limitada $X \subset \mathbb{R}^m$, a restrição $f|X$ é uniformemente contínua.

7.10. Sejam I, J intervalos da reta com interseção não-vazia. Se $f\colon I \cup J \to \mathbb{R}^n$ é tal que $f|I$ e $f|J$ são uniformemente contínuas então f é uniformemente contínua.

7.11. Considerando as seqüências de pontos $z_k = (k, 1/k)$ e $w_k = (k, 0)$ em $\mathbb{R}^2$, prove que a aplicação $\varphi\colon \mathbb{R}^2 \to \mathbb{R}$, dada por $\varphi(x, y) = xy$, não é uniformemente contínua. Use um argumento análogo para provar que uma aplicação bilinear $\varphi\colon \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^p$ só é uniformente contínua se for identicamente nula.

7.12. Sejam $f\colon X \to \mathbb{R}^n$ e $a \in X$. Suponha que, para todo $\varepsilon > 0$, exista $g\colon X \to \mathbb{R}^n$, contínua no ponto a, tal que $|f(x) - g(x)| < \varepsilon$ para todo $x \in X$. Então f é contínua no ponto a.

7.13. Considere em $\mathbb{R}^n$ a norma euclidiana. Dada uma isometria $f\colon \mathbb{R}^n \to \mathbb{R}^n$, existem um vetor $b \in \mathbb{R}^n$ e uma transformação linear (necessariamente ortogonal) $A\colon \mathbb{R}^n \to \mathbb{R}^n$ tais que $f(x) = Ax + b$ para todo $x \in \mathbb{R}^n$.

7.14. Dada uma aplicação linear $A\colon \mathbb{R}^m \to \mathbb{R}^n$, e fixadas normas em $\mathbb{R}^m$ e $\mathbb{R}^n$, a imagem por A da esfera unitária $S = \{x \in \mathbb{R}^m; |x| = 1\}$ é um conjunto limitado em $\mathbb{R}^n$. Pondo, para cada $A \in \mathcal{L}(\mathbb{R}^m;\mathbb{R}^n)$, $|A| = \sup\{|Ax|; x \in S\}$, a função $A \mapsto |A|$ é uma norma no espaço vetorial $\mathcal{L}(\mathbb{R}^m;\mathbb{R}^n)$, para a qual vale a desigualdade $|Ax| \le |A|\,|x|$ para todo $x \in \mathbb{R}^m$. Além disso, se $A \in \mathcal{L}(\mathbb{R}^m;\mathbb{R}^n)$ e $B \in \mathcal{L}(\mathbb{R}^n;\mathbb{R}^p)$ então, fixadas normas em $\mathbb{R}^m$, $\mathbb{R}^n$ e $\mathbb{R}^p$, tem-se $|BA| \le |B|\,|A|$.

§8.

8.1. O cone $C = \{(x,y,z) \in \mathbb{R}^3; z \ge 0, x^2+y^2-z^2 = 0\}$ é homeomorfo a $\mathbb{R}^2$.

8.2. Estabeleça um homeomorfismo entre $\mathbb{R}^{n+1} - \{0\}$ e $S^n \times \mathbb{R}$.

8.3. Para cada $c > 0$, o hiperbolóide de revolução $H = \{(x,y,z) \in \mathbb{R}^3; x^2+y^2-z^2 = c\}$ é homeomorfo a $S^1 \times \mathbb{R}$.

8.4. O quadrante $P = \{(x,y) \in \mathbb{R}^2; x \ge 0, y \ge 0\}$ é homeomorfo ao semi-plano superior $S = \{(x,y) \in \mathbb{R}^2; y \ge 0\}$.

8.5. Os conjuntos $X = \{(x,y) \in \mathbb{R}^2; y = 0, 0 < x < 1\}$ e $Y = \{(x,y) \in \mathbb{R}^2; y = 0\}$ são homeomorfos mas não existe um homeomorfismo $h\colon \mathbb{R}^2 \to \mathbb{R}^2$ tal que $h(X) = Y$.

8.6. Estabeleça um homeomorfismo entre os conjuntos $X = \{x \in \mathbb{R}^n; 0 < |x| \le 1\}$ (bola unitária fechada menos a origem) e $Y = \{y \in \mathbb{R}^n; |y| \ge 1\}$ (complementar da bola unitária aberta).

8.7. A "figura 8"é a reunião de dois círculos tangentes externamente em $\mathbb{R}^2$. Defina uma bijeção contínua de $\mathbb{R}$ sobre a figura 8 e mostre que sua inversa é descontínua.

8.8. Um conjunto $X \subset \mathbb{R}^n$ chama-se *topologicamente homogêneo* quando, dados $a, b \in X$ quaisquer, existe um homeomorfismo $h\colon X \to X$ tal que $h(a) = b$. Prove que, para todo $n \in \mathbb{N}$, o espaço $\mathbb{R}^n$ e a esfera S^n são topologicamente homogêneos. Por outro lado, o intervalo fechado $[0,1]$ não é topologicamente homogêneo.

8.9. Toda bola aberta em $\mathbb{R}^n$ é topologicamente homogênea.

8.10. Sejam G um grupo multiplicativo de matrizes $n \times n$ e $H \subset G$ um subgrupo. Dadas duas classes laterais aH e bH (onde $a, b \in G$), mostre que existe um homeomorfismo $\varphi\colon G \to G$ tal que $\varphi(aH) = bH$.

8.11. A projeção estereográfica $\varphi\colon S^2 - \{p\} \to \mathbb{R}^2$ transforma os círculos em S^2 que passam pelo pólo p em retas de $\mathbb{R}^2$ e os círculos em S^2 que não contêm o pólo em círculos de $\mathbb{R}^2$.

§9.

9.1. Sejam $X \subset \mathbb{R}^m$ ilimitado, $f\colon X \to \mathbb{R}^n$ uma aplicação e $a \in \mathbb{R}^n$. Diz-se que $\lim\limits_{x\to\infty} f(x) = a$ quando, para todo $\varepsilon > 0$ dado, existe $r > 0$ tal que $x \in X$, $|x| > r \Rightarrow |f(x) - a| < \varepsilon$. Prove que $\lim\limits_{x\to\infty} f(x) = a$ se, e somente se, para toda seqüência de pontos $x_k \in X$ com $\lim |x_k| = \infty$, tem-se $\lim\limits_{k\to\infty} f(x_k) = a$.

9.2. Sejam $f\colon X \to \mathbb{R}^n$ uma aplicação e a um ponto de acumulação do conjunto $X \subset \mathbb{R}^m$. Diz-se que $\lim_{x\to a} f(x) = \infty$ quando, para todo $r > 0$, existe $\delta > 0$ tal que $x \in X$, $0 < |x - a| < \delta \Rightarrow |f(x)| > r$. Prove que $\lim_{x\to a} f(x) = \infty$ se, e somente se, para toda seqüência de pontos $x_k \in X - \{a\}$, com $\lim x_k = a$, tem-se $\lim |f(x_k)| = \infty$.

9.3. Seja $f\colon X \to \mathbb{R}^n$ definida num conjunto ilimitado $X \subset \mathbb{R}^m$. Defina o que se entende por $\lim_{x\to\infty} f(x) = \infty$ e dê uma caracterização deste conceito por meio de seqüências.

9.4. Seja $p\colon \mathbb{R}^2 \to \mathbb{R}^2$ um polinômio complexo não-constante. Mostre que $\lim_{z\to\infty} p(z) = \infty$.

9.5. Seja $f\colon \mathbb{R}^2 \to \mathbb{R}$ definida por $f(x,y) = (x^2 - y)y/x^4$ se $0 < y < x^2$ e $f(x,y) = 0$ nos demais pontos. Prove que o limite de $f(x,y)$ é zero quando (x,y) tende para $(0,0)$ ao longo de qualquer reta que passe pela origem mas não se tem $\lim_{(x,y)\to(0,0)} f(x,y) = 0$.

9.6. Seja $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ definida por $f(x,y) = \dfrac{x^2 - y^2}{x^2 + y^2}$ se $x^2 + y^2 \neq 0$ e $f(0,0) = 0$. Mostre que $\lim_{x\to 0}(\lim_{y\to 0} f(x,y)) \neq \lim_{y\to 0}(\lim_{x\to 0} f(x,y))$.

9.7. Seja $h\colon B \to \mathbb{R}^n$ o homeomorfismo da bola aberta de centro 0 e raio 1 em $\mathbb{R}^n$ dado por $h(x) = x/(1-|x|)$. Fixado arbitrariamente $a \in \mathbb{R}^n$, seja $T\colon \mathbb{R}^n \to \mathbb{R}^n$ a translação $T(x) = x + a$. Considere o homeomorfismo $\varphi = h^{-1}Th\colon B \to B$. Prove que $\lim_{x\to b} \varphi(x) = b$ para todo $b \in \partial B$. Conclua que, dados arbitrariamente $c, d \in B$ existe um homeomorfismo $\bar{\varphi}\colon \overline{B} \to \overline{B}$ tal que $\bar{\varphi}(c) = d$ e $\bar{\varphi}(x) = x$ para todo $x \in \partial B$.

§10.

10.1. Um ponto $a \in X$ é aberto em X se, e somente se, é um ponto isolado de X. Conseqüentemente, um conjunto $X \subset \mathbb{R}^n$ é discreto se, e somente se, qualquer subconjunto $A \subset X$ é aberto em X.

10.2. Seja $h\colon X \to Y$ um homeomorfismo. Um conjunto $A \subset X$ é aberto em X se, e somente se, $h(A)$ é aberto em Y.

10.3. Se $A \subset \mathbb{R}^n$ é aberto então sua fronteira ∂A possui interior vazio. Dê exemplo de um conjunto $X \subset \mathbb{R}^n$ cuja fronteira ∂X seja um conjunto aberto.

10.4. Para quaisquer $X, Y \subset \mathbb{R}^n$, tem-se $\operatorname{int}(X \cap Y) = \operatorname{int} X \cap \operatorname{int} Y$ e $\operatorname{int}(X \cup Y) \supset \operatorname{int} X \cup \operatorname{int} Y$. Dê um exemplo em que esta inclusão não se reduz a uma igualdade.

10.5. Uma aplicação $f\colon X \to Y$ é aberta se, e somente se, para toda bola aberta B com centro num ponto de X, $f(B \cap X)$ é aberto em Y.

10.6. Sejam $X \subset \mathbb{R}^m$, $Y \subset \mathbb{R}^n$. Uma aplicação $f\colon X \to Y$ chama-se um *homeomorfismo local* quando, para cada ponto $x \in X$, existe A aberto em X, com $x \in A$, tal que $f|A$ é um homeomorfismo de A sobre um conjunto aberto em Y. (i) Prove que $\xi\colon \mathbb{R} \to S^1$, $\xi(t) = (\cos t, \operatorname{sen} t)$, é um homeomorfismo

local; (ii) Prove que todo homeomorfismo local é uma aplicação aberta; (iii) Se $f\colon X \to Y$ é um homeomorfismo local então, para todo $y \in Y$, a imagem inversa $f^{-1}(y)$ é um conjunto discreto.

10.7. Seja $A \subset \mathbb{R}^n$ aberto, com $n \geq 2$. Dado $a \in \mathbb{R}^n - A$, o conjunto $A \cup \{a\}$ é aberto se, e somente se, a é um ponto isolado da fronteira ∂A. Equivalentemente: existe uma bola $B = B(a;r)$ tal que $B - \{a\} \subset A$.

10.8. Seja $E \subset \mathbb{R}^n$ um subespaço vetorial. Se $E \neq \mathbb{R}^n$ então $\operatorname{int} E = \varnothing$.

10.9. O conjunto das matrizes invertíveis $n \times n$ é aberto em $\mathbb{R}^{n^2}$.

10.10. Uma aplicação linear $A\colon \mathbb{R}^n \to \mathbb{R}^n$ diz-se *positiva* quando é simétrica e, além disso, $\langle Ax, x\rangle > 0$ para todo $x \neq 0$ em $\mathbb{R}^n$. O conjunto das aplicações lineares positivas é convexo e aberto no conjunto das aplicações simétricas.

10.11. Seja G um grupo multiplicativo de matrizes $n \times n$. Se $\operatorname{int} G \neq \varnothing$ então G é aberto em $\mathbb{R}^{n^2}$.

10.12. O conjunto das aplicações lineares injetivas é aberto em $\mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$. Idem para as sobrejetivas.

10.13. Em todo conjunto aberto não vazio $U \subset \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$ existe uma aplicação linear injetiva (se $m \leq n$) ou sobrejetiva se ($m \geq n$).

10.14. Toda coleção de abertos não-vazios, dois a dois disjuntos, é enumerável.

§11.

11.1. Se um aberto A contém pontos do fecho de X então A contém pontos de X.

11.2. Dê exemplo de conjuntos $X, Y \subset \mathbb{R}^2$, tais que as projeções de X sobre os eixos são subconjuntos abertos de $\mathbb{R}$, as de Y são fechadas, mas nem X é aberto em $\mathbb{R}^2$ nem Y é fechado.

11.3. O conjunto dos pontos de acumulação de um conjunto $X \subset \mathbb{R}^n$ é fechado.

11.4. Quaisquer que sejam $X, Y \subset \mathbb{R}^n$, tem-se $\overline{X \cup Y} = \overline{X} \cup \overline{Y}$ e $\overline{X \cap Y} \subset \overline{X} \cap \overline{Y}$. Pode ocorrer $\overline{X \cap Y} \neq \overline{X} \cap \overline{Y}$.

11.5. Um conjunto $A \subset \mathbb{R}^n$ é aberto se e somente se $A \cap \overline{X} \subset \overline{A \cap X}$ para todo $X \subset \mathbb{R}^n$.

11.6. Se $X \subset \mathbb{R}^m$ e $Y \subset \mathbb{R}^n$ então $\overline{X \times Y} = \overline{X} \times \overline{Y}$ em $\mathbb{R}^{m+n}$.

11.7. $f\colon X \to \mathbb{R}^n$ é contínua se, e somente se, para todo $Y \subset X$, tem-se $f(X \cap \overline{Y}) \subset \overline{f(Y)}$.

11.8. $A \subset \mathbb{R}^n$ é aberto se, e somente se, $A \cap \overline{\mathbb{R}^n - A} = \varnothing$.

11.9. Todo subespaço vetorial $E \subset \mathbb{R}^n$ é um conjunto fechado. Se $E \neq \mathbb{R}^n$ então $\overline{\mathbb{R}^n - E} = \mathbb{R}^n$.

11.10. O fecho de um conjunto convexo é convexo.

11.11. Seja $A \subset \mathbb{R}^n$ aberto e convexo. Prove que $A = \operatorname{int} \overline{A}$. Dê exemplo de um conjunto aberto não-convexo A que seja um subconjunto próprio de $\operatorname{int} \overline{A}$.

11.12. Seja $B(X;\varepsilon)$ a reunião das bolas abertas $B(x;\varepsilon)$ de raio ε e centro em algum ponto $x \in X$. Prove que $\overline{X} = \bigcap_{\varepsilon > 0} B(X;\varepsilon)$.

11.13. Se $F \subset \mathbb{R}^n$ é fechado então sua fronteira ∂F tem interior vazio.

11.14. Um conjunto $F \subset \mathbb{R}^n$ é fechado se, e somente se, $F \supset \partial F$.

11.15. Todo conjunto fechado $F \subset \mathbb{R}^n$ é fronteira de algum conjunto $X \subset \mathbb{R}^n$.

11.16. Sejam $F \subset X \subset Y$. Se F é fechado em X e X é fechado em Y então F é fechado em Y.

11.17. Sejam F, G fechados em $X = F \cup G$. Se $f \colon X \to \mathbb{R}^n$ é tal que suas restrições $f|F$ e $f|G$ são contínuas então f é contínua.

11.18. O conjunto das matrizes invertíveis é denso em $\mathbb{R}^{n^2}$.

11.19. Sejam G um grupo multiplicativo de matrizes $n \times n$ e H um subgrupo de G. Mostre que se H for aberto em G então H será também fechado em G. (Sugestão: considere as classes laterais de H em G.)

11.20. Um conjunto $X \subset \mathbb{R}^n$ tem interior vazio se, e somente se, seu complementar é denso em $\mathbb{R}^n$.

11.21. O conjunto das matrizes $n \times n$ com determinante 1 é um fechado ilimitado com interior vazio em $\mathbb{R}^{n^2}$.

11.22. Se $A \subset X$ é aberto em X e $D \subset X$ é denso em X então $A \cap D$ é denso em A.

11.23. Dada uma seqüência de pontos $x_k \in \mathbb{R}^n$ ponhamos, para cada $k \in \mathbb{N}$, $X_k = \{x_k, x_{k+1}, \dots\}$. Prove que $\bigcap_{k \in \mathbb{N}} \overline{X}_k$ é o conjunto dos valores de aderência da seqüência (x_k) e conclua que tal conjunto é fechado.

11.24. (Teorema de Baire.) Sejam $F_1, F_2, \dots, F_i, \dots$ conjuntos fechados com interior vazio em R^n. Então $F = F_1 \cup F_2 \cup \dots$ tem interior vazio. Conclua que se $A_1, A_2, \dots, A_i, \dots$ são abertos densos em $\mathbb{R}^n$ então $A = A_1 \cap A_2 \cap \dots$ é denso em $\mathbb{R}^n$.

11.25. Todo conjunto fechado enumerável possui algum ponto isolado.

§12.

12.1. O conjunto dos valores de aderência de uma seqüência limitada é um conjunto compacto não-vazio.

12.2. As matrizes ortogonais $n \times n$ formam um subconjunto compacto de $\mathbb{R}^{n^2}$.

12.3. Todo conjunto infinito $X \subset \mathbb{R}^n$ possui um subconjunto não-compacto.

12.4. A proposição (12.4) seria falsa se tomássemos conjuntos fechados $F_1 \supset F_2 \supset \cdots \supset F_i \supset \dots$ em vez de compactos.

12.5. Seja $X \subset \mathbb{R}^{n+1} - \{0\}$ um conjunto compacto que contém exatamente um ponto em cada semi-reta de origem 0 em $\mathbb{R}^{n+1}$. Prove que X é homeomorfo à esfera unitária S^n.

12.6. Seja $X \subset \mathbb{R}^n$. Se todo conjunto homeomorfo a X for limitado então X é compacto.

12.7. Se todo conjunto $Y \subset \mathbb{R}^n$ homeomorfo a X for fechado então X é compacto.

12.8. Seja $K = [0, 2\pi] \times [0, 2\pi] \subset \mathbb{R}^2$. Defina as aplicações $f\colon K \to \mathbb{R}^3$, $g\colon K \to S^1 \times S^1 \subset \mathbb{R}^4$ e $h\colon S^1 \times S^1 \to \mathbb{R}^3$ pondo

$$f(s,t) = ((a + b\cos s)\cos t,\ (a + b\cos s)\,\text{sen}\,t, b\,\text{sen}\,s), a > b.$$
$$g(s,t) = (\cos s, \text{sen}\,s, \cos t, \text{sen}\,t),\ h(g(s,t)) = f(s,t).$$

(i) Mostre que h é bem definida e contínua.

(ii) h é um homeomorfismo de $S^1 \times S^1$ sobre $T = f(K)$ = toro gerado pela rotação de um círculo vertical de raio b e centro $(a, 0, 0)$ em torno do eixo z.

12.9. Sejam $K, L \subset \mathbb{R}^n$ compactos. O conjunto $K * L$, reunião de todos os segmentos de reta $[x, y]$ com $x \in K$, $y \in L$, é compacto. Se $F \subset \mathbb{R}^n$ é apenas fechado e $a \in \mathbb{R}^n$ então $a * F$ pode não ser fechado.

12.10. Seja $C(X)$ a interseção de todos os subconjuntos convexos de R^n que contêm $X \subset \mathbb{R}^n$. Se X é compacto então $C(X)$ também é.

12.11. Seja $X \subset \mathbb{R}^m$. Uma aplicação $f\colon X \to \mathbb{R}^n$ chama-se *localmente injetiva* quando para cada $x \in X$ existe uma bola B de centro x em $\mathbb{R}^m$ tal que $f|(B \cap X)$ é injetiva. (i) Se f é localmente injetiva então, para cada $y \in \mathbb{R}^n$, a imagem inversa $f^{-1}(y)$ é um conjunto discreto. (ii) Se f é contínua, localmente injetiva e X é compacto então, para todo $y \in \mathbb{R}^n$, a imagem inversa $f^{-1}(y)$ é um conjunto finito.

12.12. Seja $f\colon \mathbb{R}^m \to \mathbb{R}^n$ contínua. As seguintes afirmações são equivalentes:

(i) $\lim\limits_{x \to \infty} f(x) = \infty$.

(ii) A imagem inversa $f^{-1}(K)$ de todo compacto $K \subset \mathbb{R}^n$ é compacta.

Quando isto ocorre, a aplicação f diz-se *própria*. Se $f\colon \mathbb{R}^m \to \mathbb{R}^n$ é própria, então todo fechado $F \subset \mathbb{R}^m$, tem imagem $f(F)$ fechada.

12.13. Seja $X \subset \mathbb{R}^m$. Uma aplicação limitada $\varphi\colon X \to \mathbb{R}^n$ é contínua se, e somente se, seu gráfico é um subconjunto fechado de $X \times \mathbb{R}^n$.

12.14. Sejam $X \subset \mathbb{R}^m$, $K \subset \mathbb{R}^n$ compacto, $f\colon X \times K \to \mathbb{R}^p$ contínua e $c \in \mathbb{R}^p$. Suponha que, para cada $x \in X$, exista um único $y \in K$ tal que $f(x, y) = c$. Prove que esse y depende continuamente de x.

12.15. Seja $f\colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ a função contínua definida por $f(x, y) = (x^2 + y^2)(1 - xy)$. Para cada $x \in \mathbb{R}$ existe um único $y \in \mathbb{R}$ tal que $f(x, y) = 0$, mas tal y não depende continuamente de x.

12.16. Seja $f\colon \mathbb{R} \times [0, 1) \to \mathbb{R}$ definida por $f(x, y) = (x^2 + y^2)(1 - ye^{|x|})$. Para cada $x \in \mathbb{R}$, existe um único $y = \varphi(x) \in [0, 1)$ tal que $f(x, y) = 0$ mas a função $\varphi\colon \mathbb{R} \to [0, 1)$, assim definida, não é contínua.

12.17. Seja $K \subset \mathbb{R}^n$ compacto. Prove que a projeção $\pi\colon \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^m$, $\pi(x, y) = x$, transforma todo conjunto fechado $F \subset \mathbb{R}^m \times K$ num fechado $\pi(F) \subset \mathbb{R}^m$.

12.18. Sejam $K \subset \mathbb{R}^n$ compacto, $a \in \mathbb{R}^m$ e $U \subset \mathbb{R}^{m+n}$ um aberto tal que $a \times K \subset U$. Prove que existe uma bola $B \subset \mathbb{R}^m$, de centro a, tal que $B \times K \subset U$. (Sugestão: os compactos $K_i = B[a; 1/i] \times K$ têm interseção $a \times K$.)

12.19. Sejam B a bola unitária fechada de $\mathbb{R}^n$ e $a \in \mathbb{R}^n$ com $|a| < 1$. A aplicação $f\colon \mathbb{R}^n \to \mathbb{R}^n$, dada por $f(x) = (1 - |x|)a + x$, define um homeomorfismo de B sobre si mesmo, tal que $f(0) = a$ e $f(x) = x$ se $|x| = 1$. (Geometricamente, $f(x)$ é o ponto do segmento de reta $[a, x/|x|]$ tal que $[x, f(x)]$ é paralelo a $[0, a]$.) Conclua que se $B \subset \mathbb{R}^n$ é qualquer bola fechada e $a, b \in \operatorname{int} B$ então existe um homeomorfismo $h\colon B \to B$ tal que $h(a) = b$ e $h(x) = x$ para todo $x \in \partial B$. (Outra demonstração deste fato é sugerida no Exercício 9.7.)

12.20. Toda aplicação localmente Lipschitziana definida num conjunto compacto é Lipschitziana. (Vide Exercício 7.6.)

§13.

13.1. Se $U \subset R^n$ é um aberto limitado, não existem $x_0, y_0 \in U$ tais que $|x_0 - y_0| = \operatorname{diam} U$.

13.2. Seja $B = B[a; r] \subset \mathbb{R}^n$. Para todo $x \in \mathbb{R}^n$, tem-se $d(x, B) = \max\{0, |x-a|-r\}$.

13.3. Seja $T = \mathbb{R}^n - B[a.r]$. Para todo $x \in \mathbb{R}^n$, tem-se $d(x, T) = \max\{0, r - |x - a|\}$.

13.4. $d(S, T) = \inf\limits_{s \in S} d(x, T)$.

13.5. A função de Urysohn de um par de fechados disjuntos $F, G \subset \mathbb{R}^n$ é uniformemente contínua se, e somente se, $d(F, G) > 0$.

13.6. Considerando em $\mathbb{R}^n$ a norma euclidiana, sejam $F \subset \mathbb{R}^n$ um conjunto fechado convexo, a um ponto de $\mathbb{R}^n$ e $y_0 \in F$ tal que $|a - y_0| = d(a, F)$. Mostre que, para todo $x \in F$ tem-se $\langle x - y_0, a - y_0 \rangle \leq 0$.

§14.

14.1. Uma decomposição $X = A \cup B$ é uma cisão se, e somente se, nenhum dos conjuntos A, B contêm um ponto aderente ao outro. Isto se exprime por $(\overline{A} \cap B) \cup (A \cap \overline{B}) = \varnothing$.

14.2. Um subconjunto conexo não-vazio $X \subset \mathbb{Q}^n$ consta de um único ponto.

14.3. Seja $E \subset \mathbb{R}^n$ um subespaço vetorial próprio. O complementar $\mathbb{R}^n - E$ é conexo se, e somente se, $\dim E \leq n - 2$.

14.4. O conjunto das matrizes invertíveis $n \times n$ é um aberto desconexo em $\mathbb{R}^{n^2}$. Também é desconexo (mas não aberto) o conjunto das matrizes ortogonais.

14.5. Se $X \subset \mathbb{R}^m$ é compacto então toda aplicação contínua aberta $f\colon X \to S^n$ é sobrejetiva.

14.6. Seja $X \subset \mathbb{R}^m$. Uma aplicação $f\colon X \to \mathbb{R}^n$ diz-se *localmente constante* quando para cada $x \in X$ existe uma bola B de centro x tal que $f|(B \cap X)$ é constante. X é conexo se, e somente se, toda aplicação localmente constante $f\colon X \to \mathbb{R}^n$ é constante.

14.7. Toda aplicação contínua $f\colon X \to \mathbb{R}^n$ cuja imagem $f(X)$ é um conjunto discreto é localmente constante.

14.8. Toda aplicação localmente constante $f\colon X \to \mathbb{R}^n$ tem imagem enumerável.

14.9. Um conjunto conexo enumerável $X \subset \mathbb{R}^n$ possui no máximo um ponto.

14.10. Se $X \subset \mathbb{R}^m$ é conexo por caminhos e $f\colon X \to \mathbb{R}^n$ é contínua então $f(X)$ é conexo por caminhos.

14.11. Se $X \subset \mathbb{R}^m$, $Y \subset \mathbb{R}^n$ são conexos por caminhos então $X \times Y \subset \mathbb{R}^{m+n}$ é conexo por caminhos.

14.12. A reunião de uma família de conjuntos conexos por caminhos com um ponto em comum é conexa por caminhos.

14.13. O fecho de um conjunto conexo por caminhos pode não ser conexo por caminhos.

14.14. Seja B uma bola (fechada ou aberta) em $\mathbb{R}^n$, com $n \geq 2$. Para todo $x \in B$, o conjunto $B - \{x\}$ é conexo.

14.15. Seja $B \subset \mathbb{R}^n$ uma bola fechada na norma euclidiana. Para todo subconjunto $X \subset \partial B$, $B - X$ é convexo. Numa norma arbitrária, $B - X$ é conexo mas não necessariamente convexo.

14.16. O conjunto das matrizes ortogonais $n \times n$ com determinante positivo (isto é, igual a 1) é conexo por caminhos. Conclua que o conjunto de todas as matrizes ortogonais tem duas componentes conexas. Resultados análogos valem para as matrizes invertíveis $n \times n$.

14.17. As componentes conexas de um subconjunto aberto em $\mathbb{R}^n$ são conjuntos abertos.

14.18. Um conjunto $M \subset \mathbb{R}^n$ chama-se uma *superfície de dimensão m* quando, para todo $x \in M$, existem A aberto em M contendo x, A_0 aberto em $\mathbb{R}^m$ e um homeomorfismo $\varphi\colon A_0 \to A$. Prove que a esfera $S^n \subset \mathbb{R}^{n+1}$ é uma superfície de dimensão n.

14.19. Uma superfície é conexa se, e somente se, é conexa por caminhos.

14.20. Se $M \subset \mathbb{R}^k$ é uma superfície de dimensão m e $N \subset \mathbb{R}^p$ é uma superfície de dimensão n então $M \times N \subset \mathbb{R}^{k+p}$ é uma superfície de dimensão $m+n$. Em particular, o toro é uma superfície de dimensão 2.

14.21. Seja M uma superfície conexa. Dados dois pontos arbitrários $a, b \in M$, existe um homeomorfismo $h\colon M \to M$ tal que $h(a) = b$. (Fixe $a \in M$ e seja X o conjunto dos pontos $x \in M$ tais que existe um homeomorfismo $h\colon M \to M$ com $h(a) = x$. Para provar que X e $M - X$ são ambos abertos, use o Exercício 12.19 ou 9.7.)

§15.

15.1. Dadas $A, B \in \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$, ponha $\langle A, B\rangle = \operatorname{tr} \cdot AB^*$(= traço da matriz AB^*). Prove que $\langle A, B\rangle = \sum_{i,j} a_{ij} b_{ij}$, conclua que $\langle A, B\rangle = \operatorname{tr} \cdot A^*B = \operatorname{tr} \cdot B^*A = \operatorname{tr} \cdot BA^*$, que $\langle A, B\rangle$ é um produto interno em $\mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$ e que, pondo $||A|| = \sqrt{\langle A, A\rangle}$, tem-se $|Ax| \leq ||A|| \cdot |x|$ para todo $x \in \mathbb{R}^m$ (onde $|x|$ e $|Ax|$ são normas euclidianas).

15.2. Seja G o grupo das matrizes invertíveis $n \times n$. Mostre que se $A \in G$ e $|Ax| \geq |c|\,|x|$ para todo $x \in \mathbb{R}^n$ então $|A^{-1}| \leq 1/c$. Conclua que se $X \in G$ e $|X - A| < c/2$ então $|X^{-1}| \leq 2/c$. Em seguida, use a identidade $X^{-1} - A^{-1} =$

$X^{-1}(I - XA^{-1})$ para mostrar que $\lim_{X \to A} X^{-1} = A^{-1}$. Logo $f \colon G \to G$, dada por $f(X) = X^{-1}$, é contínua.

15.3. Dada $A \in \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$, supomos fixadas uma norma em $\mathbb{R}^m$, outra em $\mathbb{R}^n$ e pomos, como no texto, $|A| = \sup\{|Ax|; x \in \mathbb{R}^m, |x| = 1\}$. Prove que $|A| = \inf\{c \in \mathbb{R}; |Ax| \leq c|x| \text{ para todo } x \in \mathbb{R}^m\}$.

15.4. Nas condições do exercício anterior, se tomarmos em $\mathbb{R}^m$ e $\mathbb{R}^n$ as normas provenientes do produto interno, prove que $|A| \leq \sqrt{\sum_{i,j} a_{ij}^2}$, onde (a_{ij}) é a matriz de A.

Capítulo II

Caminhos no Espaço Euclidiano

Nota. A fim de não restringir desnecessariamente a generalidade no estudo da integração e do comprimento de caminhos, divergiremos, neste capítulo, da terminologia adotada no anterior, onde todos os caminhos eram contínuos, por definição. A partir do próximo capítulo, porém, os caminhos voltarão a ser contínuos.

1 Caminhos diferenciáveis

Um *caminho* em $\mathbb{R}^n$ é uma aplicação $f\colon I \to \mathbb{R}^n$, cujo domínio é um intervalo $I \subset \mathbb{R}$. Para cada $t \in I$, temos $f(t) = (f_1(t), \ldots, f_n(t))$. As n funções $f_i\colon I \to \mathbb{R}$ são chamadas as *funções coordenadas* de f. Escreve-se, então, $f = (f_1, \ldots, f_n)$.

O caminho $f\colon I \to \mathbb{R}^n$ é contínuo no ponto $a \in I$ se, e somente se, cada uma das suas funções coordenadas é contínua nesse ponto. Mais geralmente, sabemos que se $f = (f_1, \ldots, f_n)$ é definida num conjunto $X \subset \mathbb{R}$ e a é um ponto de acumulação de X, então $\lim\limits_{t\to a} f(t) = b = (b_1, \ldots, b_n)$ se, e somente se, $\lim\limits_{t\to a} f_i(t) = b_i$ para cada $i = 1, 2, \ldots, n$. Como $X \subset \mathbb{R}$, têm sentido os limites laterais $\lim\limits_{t\to a^-} f(t)$ e $\lim\limits_{t\to a^+} f(t)$, o primeiro quando a é ponto de acumulação à esquerda e o segundo quando a é ponto de acumulação à direita.

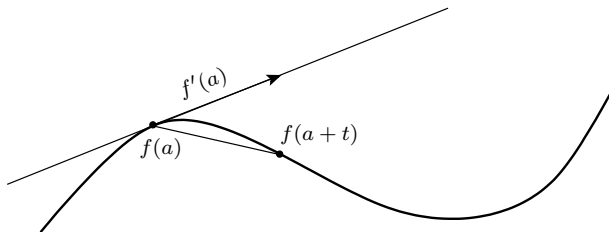
O *vetor velocidade* do caminho $f\colon I \to \mathbb{R}^n$ no ponto $a \in I$ é, por

definição, o limite

$$f'(a) = \lim_{t\to 0} \frac{f(a+t) - f(a)}{t},$$

quando tal limite existe. Às vezes escreveremos $\dfrac{df}{dt}(a)$ ou $Df(a)$ em vez de $f'(a)$. A norma $|f'(a)|$ chama-se a *velocidade escalar* de f no ponto a.

Quando o caminho f possui vetor velocidade num ponto, dizemos que f é *diferenciável* nesse ponto. Se existe $f'(a)$ para todo $a \in I$, dizemos que $f\colon I \to \mathbb{R}^n$ é um *caminho diferenciável*.

O vetor velocidade $f'(a)$, quando é diferente de zero, determina a *reta tangente* ao caminho f no ponto $f(a)$, a saber, a reta $L = \{f(a) + t \cdot f'(a); t \in \mathbb{R}\}$.

Como as coordenadas do "vetor secante" $[f(a+t) - f(a)]/t$ são os números $[f_i(a+t)-f_i(a)]/t$, vemos que o caminho f possui vetor tangente no ponto a se, e somente se, cada uma das suas funções coordenadas f_i possui derivada nesse ponto. No caso afirmativo, tem-se $f'(a) = (f_1'(a), \ldots, f_n'(a))$. Assim, o caminho $f\colon I \to \mathbb{R}^n$ é diferenciável se, e somente se, cada uma das suas funções coordenadas $f_i\colon I \to \mathbb{R}$ é uma função real diferenciável no intervalo I.

Em particular, a diferenciabilidade do caminho f não depende da norma que estivermos utilizando em $\mathbb{R}^n$.

A fim de que o caminho $f\colon I \to \mathbb{R}^n$ seja diferenciável no ponto $a \in I$, é necessário e suficiente que exista um vetor $v \in \mathbb{R}^n$ tal que, para $a+t \in I$ se tenha

$$f(a+t) = f(a) + t \cdot v + r(t), \text{ onde } \lim_{t\to 0} \frac{r(t)}{t} = 0.$$

No caso afirmativo, tem-se $v = f'(a)$.

Com efeito, a primeira igualdade acima significa que, para $t \neq 0$,

$$\frac{r(t)}{t} = \frac{f(a+t) - f(a)}{t} - v.$$

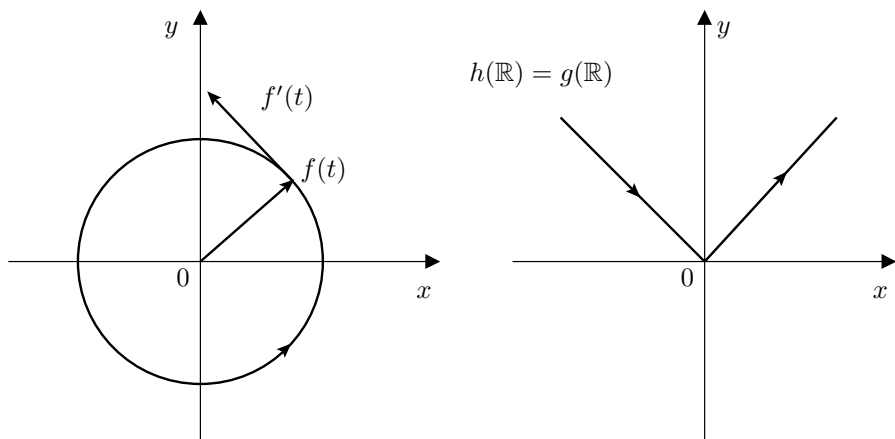
A condição de diferenciabilidade de f no ponto a pode ser expressa, equivalentemente, assim:

$$f(a+t) = f(a) + t[f'(a) + \rho(t)], \text{ onde } \lim_{t \to 0} \rho(t) = 0.$$

Basta pôr $\rho(t) = r(t)/t$, se $t \neq 0$, e $\rho(0) = 0$.

A noção de derivada lateral se define, para caminhos, de modo análogo ao caso de funções reais. Tem-se $f'(a+) = (f_1'(a+), \ldots, f_n'(a+))$ e $f'(a-) = (f_1'(a-), \ldots, f_n'(a-))$. Existe o vetor velocidade de f no ponto a se, e somente se, existem e são iguais as derivadas laterais nesse ponto.

Exemplo 1. Seja $f\colon \mathbb{R} \to \mathbb{R}^2$ dado por $f(t) = (\cos t, \operatorname{sen} t) = e^{it}$. A imagem do caminho f é o círculo unitário S^1. Para todo $t \in \mathbb{R}$, o vetor velocidade de f é $f'(t) = (-\operatorname{sen} t, \cos t) = ie^{it}$. A velocidade escalar é constante: $|f'(t)| = 1$.

Exemplo 2. O caminho $g\colon \mathbb{R} \to \mathbb{R}^2$, definido por $g(t) = (t, |t|)$, possui vetor velocidade $g'(t) = (1, 1)$ para todo $t > 0$ e $g'(t) = (1, -1)$ para $t < 0$. No ponto $t = 0$, as derivadas laterais $g'(0-) = (1, -1)$ e $g'(0+) = (1, 1)$ existem mas são diferentes. Logo g não possui vetor velocidade no ponto 0. A imagem de g é o gráfico da função $y = |x|$, que apresenta um ponto anguloso na origem. Podemos descrever a mesma imagem por meio de outra "parametrização", considerando, por exemplo, o caminho $h\colon \mathbb{R} \to \mathbb{R}^2$, com $h(t) = (t^3, |t|t^2)$. Temos $h(\mathbb{R}) = g(\mathbb{R})$. Para $t < 0$, $h'(t) = (3t^2, -3t^2)$ e, para $t > 0$, $h'(t) = (3t^2, 3t^2)$. No ponto $t = 0$, as derivadas laterais são ambas nulas, logo existe $h'(0) = (0, 0)$. Em outras palavras: para descrever a rota $h(\mathbb{R})$, o ponto cuja posição no tempo t é $h(t)$ precisou dar uma parada instantânea ao atingir o ponto anguloso $(0, 0)$ de sua trajetória. (Veja Exercício 1.15.)

As propriedades usuais da derivada de uma função real de uma variável real, quando aplicadas às funções coordenadas de um caminho, conduzem imediatamente às seguintes regras de derivação:

$$\frac{d}{dt}\,[f(t)+g(t)] = f'(t)+g'(t),$$
$$\frac{d}{dt}\,[\alpha(t)\cdot f(t)] = \alpha'(t)f(t)+\alpha(t)f'(t),$$
$$\frac{d}{dt}\,\langle f(t),g(t)\rangle = \langle f'(t),g(t)\rangle + \langle f(t),g'(t)\rangle,$$
$$\frac{d}{dt}\,|f(t)| = \frac{\langle f(t),f'(t)\rangle}{|f(t)|} \text{ se } f(t)\neq 0.$$

Acima, $f, g\colon I \to \mathbb{R}^n$ são caminhos diferenciáveis e $\alpha\colon I \to \mathbb{R}$ é uma função diferenciável. A norma $|f(t)|$ é a que provém do produto interno $\langle f(t), g(t)\rangle$, isto é, $|f(t)| = \sqrt{\langle f(t), f(t)\rangle}$. Esta hipótese será feita sempre que tomarmos a derivada de $|f(t)|$.

Exemplo 3. Se a norma não provém de um produto interno, podemos perfeitamente ter um caminho diferenciável $f\colon I \to \mathbb{R}^n$, com $f(t)\neq 0$ para todo $t\in I$, para o qual a função $\varphi(t) = |f(t)|$ não é diferenciável. Por exemplo, consideremos a norma do máximo em $\mathbb{R}^2$. O caminho $f\colon \mathbb{R}\to\mathbb{R}^2$, dado por $f(t) = (1,t)$ é obviamente diferenciável. Mas $|f(t)| = |t|$ se $|t|\geq 1$ e $|f(t)| = 1$ se $|t|\leq 1$. Logo, a função $|f(t)|$ não possui derivada em cada um dos pontos $t=-1$ e $t=1$.

Resulta da fórmula da derivada de $|f(t)|$ que, dado um caminho diferenciável $f\colon I\to\mathbb{R}^n$, o vetor $f(t)$ tem comprimento constante se, e

somente se, para cada valor do parâmetro t, o vetor velocidade $f'(t)$ é perpendicular ao "vetor posição" $f(t)$.

Exemplo 4. O caminho $f\colon \mathbb{R} \to \mathbb{R}^2$, dado por $f(t) = (\cos t, \operatorname{sen} t)$, é tal que $|f(t)| = 1$ para todo t. Por isso $f'(t) = (-\operatorname{sen} t, \cos t)$ é, para todo $t \in \mathbb{R}$, perpendicular a $f(t)$. Neste exemplo, temos também $|f'(t)| = 1$ mas isto é acidental. Se tomarmos $g\colon \mathbb{R} \to \mathbb{R}^2$, com $g(t) = (\cos t^2, \operatorname{sen} t^2)$, ainda vale $|g(t)| = 1$ para todo t, mas $g'(t) = (-2t \operatorname{sen} t^2, 2t \cos t^2)$ tem comprimento variável, igual a $2|t|$. Continua porém sendo $g'(t)$ perpendicular a $g(t)$.

Se o caminho $f\colon I \to \mathbb{R}^n$ é diferenciável, tem sentido considerar o caminho derivado $f'\colon I \to \mathbb{R}^n$ e indagar se ele é contínuo, diferenciável, etc. Quando f' é contínuo, diz-se que f é um *caminho de classe* C^1. Pode-se então investigar a existência da derivada de f'. Quando existe, o vetor $(f')'(a) = f''(a)$ chama-se a *derivada segunda* de f no ponto a, ou o *vetor aceleração* do caminho f no ponto a. Tem-se $f''(a) = (f_1''(a), \ldots, f_n''(a))$. Se existe $f''(t)$ para todo $t \in I$, diz-se que f é *duas vezes diferenciável* e fica definido o caminho $f''\colon I \to \mathbb{R}^n$. Quando f'' é contínuo, dizemos que f é um *caminho de classe* C^2.

Prosseguindo, diremos que o caminho $f\colon I \to \mathbb{R}^n$ é $p+1$ vezes diferenciável quando existir o caminho $f^{(p)}\colon I \to \mathbb{R}^n$ (derivada de ordem p de f) e for derivável. Então poremos $f^{(p+1)} = (f^{(p)})'$. Quando $f^{(p)}$ for de classe C^1, diremos que f é de classe C^{p+1}.

Diremos que o caminho $f\colon I \to \mathbb{R}^n$ é *p vezes diferenciável no ponto* $a \in I$ quando existir $\delta > 0$ tal que f é de classe C^{p-1} no intervalo $J = \{t \in I; |t-a| < \delta\}$ e $f^{(p-1)}$ for diferenciável no ponto a.

Por extensão, diremos que um caminho contínuo é de *classe* C^0 e que $f = f^{(0)}$ é sua própria "derivada de ordem zero".

Quando existirem as derivadas de todas as ordens do caminho f, diremos que f é infinitamente derivável, ou *de classe* C^∞.

Para $0 \le p \le \infty$, escreveremos $f \in C^p$ para indicar que f é um caminho de classe C^p. Dado $f = (f_1, \ldots, f_n)$, tem-se $f \in C^p$ se, e somente se, $f_i \in C^p$ para cada $i = 1, \ldots, n$.

Seja $p > 0$. Diremos que o caminho $f\colon I \to \mathbb{R}^n$ é *de classe* C^p *por partes* quando f for contínuo e, além disso, possuir derivadas contínuas até a ordem p, inclusive, salvo num conjunto finito de pontos de I. Nesses pontos, f deve possuir derivadas laterais contínuas até a ordem p, inclusive.

Exemplo 5. Para todo $p \geq 0$, o caminho $f\colon \mathbb{R} \to \mathbb{R}^2$, definido por $f(t) = (t^{p+1}, t^p|t|)$, é de classe C^p, e de classe C^∞ por partes.

2 Integral de um caminho

Seja $f\colon [a,b] \to \mathbb{R}^n$ um caminho limitado, definido no intervalo compacto $[a,b]$. Sabemos que uma *partição* de $[a,b]$ é um conjunto finito $P = \{t_0 < t_1 < \cdots < t_k\}$ onde $t_0 = a$ e $t_k = b$. A *norma* da partição P é o número $|P| = \max\limits_{1\leq i\leq k}(t_i - t_{i-1})$. Uma *partição pontilhada* é um par $P^* = (P, \xi)$ onde P é uma partição e $\xi = (\xi_1, \ldots, \xi_k)$ é tal que $t_{i-1} \leq \xi_i < t_i$ para todo $i = 1, \ldots, k$. Dados f e uma partição pontilhada P^*, formamos a *soma de Riemann*:

$$\Sigma(f; P^*) = \sum_{i=1}^{k}(t_i - t_{i-1})f(\xi_i).$$

Diremos que o vetor $v \in \mathbb{R}^n$ é o *limite* das somas de Riemann $\Sigma(f; P^*)$ quando a norma de P tende para zero se, para todo $\varepsilon > 0$ dado, existir $\delta > 0$ tal que $|P| < \delta \Rightarrow |v - \Sigma(f; P^*)| < \varepsilon$. Neste caso, escreveremos $v = \lim\limits_{|P|\to 0} \Sigma(f; P^*)$, diremos que o caminho f é *integrável*, chamaremos v a *integral* do caminho f no intervalo $[a,b]$ e usaremos a notação $v = \int_a^b f(t)\,dt$. Portanto

$$\int_a^b f(t)\,dt = \lim_{|P|\to 0} \Sigma(f; P^*).$$

Nem sempre, é claro, existe o limite acima. A fim de que o caminho limitado $f = (f_1, \ldots, f_n)\colon [a,b] \to \mathbb{R}^n$ seja integrável, é necessário e suficiente que cada uma das funções coordenadas $f_i\colon [a,b] \to \mathbb{R}$ seja integrável. No caso afirmativo, temos

$$\int_a^b f(t)\,dt = \left(\int_a^b f_1(t)\,dt, \ldots, \int_a^b f_n(t)\,dt\right).$$

Da igualdade acima, e da propriedade análoga para funções reais, resulta que se $f\colon [a,b] \to \mathbb{R}^n$ é integrável e $c \in [a,b]$ então $f|[a,c]$ e $f|[c,b]$ são integráveis e

$$\int_a^b f(t)\,dt = \int_a^c f(t)\,dt + \int_c^b f(t)\,dt.$$

Seja D o conjunto dos pontos de descontinuidade do caminho limitado $f\colon [a,b] \to \mathbb{R}^n$. Para cada $i = 1, \ldots, n$, seja D_i o conjunto dos pontos de descontinuidade da sua i-ésima função coordenada $f_i\colon [a,b] \to \mathbb{R}$. Temos $D = D_1 \cup \cdots \cup D_n$. Sabemos que f_i é integrável se, e somente se, D_i tem medida nula. Assim, o caminho f é integrável se, e somente se, o conjunto D dos seus pontos de descontinuidade tem medida nula.

A integrabilidade de f e o valor $\int_a^b f(x)\,dx$ não dependem da norma tomada em $\mathbb{R}^n$.

Exemplo 6. Seja $f\colon [0, 2\pi] \to \mathbb{R}^2$, $f(t) = (\cos t, \operatorname{sen} t)$. Então

$$\int_0^{2\pi} f(t)\,dt = \left(\int_0^{2\pi} \cos t\,dt, \int_0^{2\pi} \operatorname{sen} t\,dt\right) = (0,0).$$

Por outro lado, se $g\colon [0,1] \to \mathbb{R}^2$ é dado por $g(t) = (t, t^2)$, então $\int_0^1 g(t)\,dt = (1/2, 1/3)$.

Sejam $f, g\colon [a,b] \to \mathbb{R}^n$ caminhos integráveis. Segue-se da definição que, para $\alpha, \beta \in \mathbb{R}$ arbitrários, o caminho $\alpha f + \beta g$ é integrável e

$$\int_a^b [\alpha f(t) + \beta g(t)]\,dt = \alpha \int_a^b f(t)\,dt + \beta \int_a^b g(t)\,dt.$$

A condição de integrabilidade em termos da medida do conjunto dos pontos de descontinuidade mostra que $f\colon [a,b] \to \mathbb{R}^n$ integrável implica $|f|\colon [a,b] \to \mathbb{R}$ integrável. Além disso, para qualquer norma tomada em $\mathbb{R}^n$, temos:

$$\left|\int_a^b f(t)\,dt\right| \le \int_a^b |f(t)|\,dt.$$

Para verificar esta desigualdade, basta observar que, dada qualquer partição pontilhada $P^* = (P, \xi)$, tem-se

$$|\Sigma(f; P^*)| = |\Sigma(t_i - t_{i-1}) f(\xi_i)| \le \Sigma(t_i - t_{i-1}) |f(\xi_i)| = \Sigma(|f|; P^*).$$

Logo

$$\left|\int_a^b f(t)\,dt\right| = |\lim_{|P|\to 0} \Sigma(f; P^*)| = \lim_{|P|\to 0} |\Sigma(f; P^*)| \le \lim_{|P|\to 0} \Sigma(|f|; P^*) =$$
$$= \int_a^b |f(t)|\,dt.$$

Daí resulta que se $|f(t)| \leq M$ para todo $t \in [a, b]$ então $|\int_a^b f(t)\,dt| \leq M(b-a)$. Deve-se notar, porém, que, para $n > 1$, se $f\colon [a, b] \to \mathbb{R}^n$ é contínuo e $|f(t)| \geq c > 0$ para todo $t \in [a, b]$, não se pode concluir que $\int_a^b f(t)\,dt \neq 0$. (Veja o Exemplo 6 acima.)

Diretamente a partir da definição vemos que se $f\colon [a, b] \to \mathbb{R}^m$ é integrável e $A\colon \mathbb{R}^m \to \mathbb{R}^n$ é uma transformação linear, então $A \cdot f\colon [a, b] \to \mathbb{R}^n$ é integrável e

$$\int_a^b A \cdot f(t)\,dt = A \cdot \left(\int_a^b f(t)\,dt \right).$$

3 Os teoremas clássicos do Cálculo

Regra da Cadeia. *Sejam $\varphi\colon I \to J$ uma função real diferenciável no ponto $a \in I$ e $f\colon J \to \mathbb{R}^n$ um caminho diferenciável no ponto $b = \varphi(a)$. Então $f \circ \varphi\colon I \to \mathbb{R}^n$ é um caminho diferenciável no ponto a e, além disso, $(f \circ \varphi)'(a) = \varphi'(a) \cdot f'(\varphi(a))$.*

Basta aplicar a regra da cadeia às funções $f_i \circ \varphi$, que são as funções coordenadas do caminho $f \circ \varphi$.

Note que, na fórmula acima, $\varphi'(a)$ é um número e $f'(\varphi(a))$ é um vetor. Pondo $x = \varphi(t)$ e adotando as notações clássicas $\dfrac{df}{dt} = (f \circ \varphi)'$, $\dfrac{df}{dx} = f'$, a regra da cadeia se lê: $\dfrac{df}{dt} = \dfrac{dx}{dt} \cdot \dfrac{df}{dx} \cdot$

Podemos interpretar a função composta $t \mapsto f(\varphi(t))$ como uma mudança de variável no caminho f. A fórmula $(f \circ \varphi)'(a) = \varphi'(a) \cdot f'(\varphi(a))$ diz que, ao efetuar-se uma tal mudança de variável no caminho f, (o que equivale a descrever o mesmo percurso de outra maneira) o vetor velocidade em cada ponto se altera apenas pela multiplicação por um escalar.

Exemplo 7. Sejam $f\colon \mathbb{R} \to \mathbb{R}^2$ e $\varphi\colon \mathbb{R} \to \mathbb{R}$ definidas por $f(x) = (\cos x, \operatorname{sen} x)$ e $\varphi(t) = t^2$. Então $f \circ \varphi\colon \mathbb{R} \to \mathbb{R}^2$ é o caminho definido por $f \circ \varphi(t) = (\cos t^2, \operatorname{sen} t^2)$. Para todo $t \in \mathbb{R}$, temos $(f \circ \varphi)'(t) = (-2t \operatorname{sen} t^2, 2t \cos t^2) = 2t \cdot f'(\varphi(t))$, múltiplo escalar do vetor velocidade de f no ponto $\varphi(t)$.

Do mesmo modo que a regra da cadeia, os seis primeiros teoremas abaixo se demonstram simplesmente observando que se $f = (f_1, \ldots, f_n)$ então $f' = (f_1', \ldots, f_n')$ e $\int_a^b f(t)\,dt = \left(\int_a^b f_1(t)\,dt, \ldots, \int_a^b f_n(t)\,dt \right)$ e,

em seguida, aplicando o teorema correspondente para funções reais de uma variável real, demonstrado no vol. 1.

Mudança de variável na integral. *Sejam* $f\colon [a,b] \to \mathbb{R}^n$ *um caminho contínuo e* $\varphi\colon [c,d] \to [a,b]$ *uma função com derivada integrável. Então*

$$\int_{\varphi(c)}^{\varphi(d)} f(x)\,dx = \int_c^d f(\varphi(t)) \cdot \varphi'(t)\,dt.$$

Teorema Fundamental do Cálculo. *Seja* $f\colon [a,a+h] \to \mathbb{R}^n$ *um caminho com derivada integrável. Então*

$$f(a+h) - f(a) = \int_a^{a+h} f'(t)\,dt = h\int_0^1 f'(a+th)\,dt.$$

Seja $f\colon [a,b] \to \mathbb{R}^n$ um caminho integrável. A *integral indefinida* de f é o caminho $F\colon [a,b] \to \mathbb{R}^n$, definido por $F(x) = \int_a^x f(t)\,dt$. Como vimos acima, se $|f(t)| \leq M$ para todo $t \in [a,b]$, então $|F(x) - F(y)| \leq M|x-y|$ sejam quais forem $x, y \in [a,b]$, logo F satisfaz a uma condição de Lipschitz.

Derivação da integral indefinida. *Se o caminho integrável* $f\colon [a,b] \to \mathbb{R}^n$ *é contínuo no ponto* c*, então* F *é diferenciável nesse ponto e* $F'(c) = f(c)$.

Fórmula de Taylor infinitesimal. *Seja* $f\colon I \to \mathbb{R}^n$ *um caminho* p *vezes diferenciável no ponto* a*. Para todo* h *tal que* $a+h \in I$*, escrevamos:*

$$f(a+h) = f(a) + h \cdot f'(a) + \cdots + \frac{h^p}{p!}\, f^{(p)}(a) + r_p(h).$$

Então $\lim\limits_{h\to 0} \dfrac{r_p(h)}{h^p} = 0.$

Fórmula de Taylor com resto integral. *Seja* $f\colon [a,a+h] \to \mathbb{R}^n$ *um caminho* p *vezes diferenciável no intervalo* $[a,a+h]$*, com* $f^{(p)}$ *integrável. Então*

$$f(a+h) = f(a) + h \cdot f'(a) + \cdots + \frac{h^{p-1}}{(p-1)!}\, f^{(p-1)}(a) + r_p$$

onde

$$r_p = \frac{h^p}{(p-1)!}\int_0^1 (1-t)^{p-1}\, f^{(p)}(a+th)dt =$$

$$= \frac{1}{(p-1)!}\int_a^{a+h} (a+h-x)^{p-1} \cdot f^{(p)}(x)\,dx.$$

Diferenciabilidade uniforme. *Um caminho* $f\colon [a,b] \to \mathbb{R}^n$ *é de classe* C^1 *se, e somente se, é uniformemente diferenciável.* (*Isto é, para todo* $\varepsilon > 0$, *existe* $\delta > 0$ *tal que*

$$x, x+h \in [a,b],\ 0 < |h| < \delta \Rightarrow |f(x+h) - f(x) - f'(x) \cdot h| < \varepsilon |h|.)$$

Isto decorre do teorema análogo para funções reais (Vol. 1, pág. 277), de acordo com as seguintes observações:

1º A diferenciabilidade uniforme de f não depende da norma considerada em $\mathbb{R}^n$ (levando em conta que duas normas quaisquer são equivalentes);

2º O caminho f é uniformemente diferenciável na norma do máximo se, e somente se, cada uma das suas funções coordenadas f_i o é;

3º Uma função $f_i\colon [a,b] \to \mathbb{R}$ é uniformemente diferenciável se, e somente se, é de classe C^1.

Examinaremos agora o Teorema do Valor Médio para caminhos. De início observemos que, dado um caminho diferenciável $f\colon [a,b] \to \mathbb{R}^n$, com $n > 1$, nem sempre existe $c \in [a,b]$ tal que $f(b) - f(a) = (b-a) \cdot f'(c)$. Basta tomar $f\colon [0, 2\pi] \to \mathbb{R}^2$, $f(t) = (\cos t, \operatorname{sen} t)$. Como $|f'(t)| = 1$ para todo $t \in [0, 2\pi]$ e $f(2\pi) - f(0) = 0$, não pode existir $c \in [0, 2\pi]$ com $f(2\pi) - f(0) = 2\pi \cdot f'(c)$. Assim, para caminhos num espaço euclidiano de dimensão maior do que 1, não vale um teorema do valor médio sob forma de igualdade. Tem-se, porém, uma desigualdade muito útil, que demonstraremos agora.

Teorema do valor médio. *Seja* $f\colon [a,b] \to \mathbb{R}^n$ *um caminho contínuo, diferenciável no intervalo aberto* (a,b). *Se* $|f'(t)| \leq M$ *para todo* $t \in (a,b)$, *então* $|f(b) - f(a)| \leq M \cdot (b-a)$.

O método de tomar cada função coordenada e aplicar a ela o teorema análogo para funções reais não funciona aqui. Vamos apresentar duas demonstrações desta importante desigualdade.

Primeira demonstração: (Válida apenas quando f' é integrável em cada subintervalo compacto $[c,d] \subset (a,b)$.) Pelo Teorema Fundamental do Cálculo, temos

$$f(d) - f(c) = \int_c^d f'(t)\, dt, \quad \text{donde} \quad |f(d) - f(c)| \leq M \cdot (d-c)$$

para todo $[c, d] \subset (a, b)$. Como f é contínua nos pontos a e b, fazendo $c \to a$ e $d \to b$ na última desigualdade acima, obtemos $|f(b) - f(a)| \leq M \cdot (b - a)$.

Segunda demonstração: (Válida em geral.) – Será baseada em dois lemas:

Lema 1. *Seja $f\colon I \to \mathbb{R}^n$ diferenciável no ponto $c \in I$. Dadas as seqüências de números $a_k \neq b_k$ em I, com $a_k \leq c \leq b_k$ e $\lim a_k = \lim b_k = c$, tem-se*

$$f'(c) = \lim_{k\to\infty} \frac{f(b_k) - f(a_k)}{b_k - a_k} \cdot$$

Demonstração: Não há perda de generalidade em supor que $a_k < c < b_k$ para todo k pois, se existirem infinitos índices k para os quais $a_k = c$ ou $b_k = c$, eles determinarão subseqüências de $[f(b_k) - f(a_k)]/(b_k - a_k)$ que obviamente convergirão para $f'(c)$. Ora, podemos escrever

$$(*) \qquad \frac{f(b_k) - f(a_k)}{b_k - a_k} = (1 - t_k)\,\frac{f(b_k) - f(c)}{b_k - c} + t_k\,\frac{f(a_k) - f(c)}{a_k - c}\,,$$

onde $t_k = (a_k - c)/(a_k - b_k)$ e portanto $1 - t_k = (b_k - c)/(b_k - a_k)$. Note que $0 \leq t_k \leq 1$. Se chamarmos de Q_k, R_k e S_k os três quocientes da igualdade (*), na ordem em que aparecem, temos $Q_k = (1 - t_k)R_k + t_k S_k$. Logo, para cada $k \in \mathbb{N}$, o ponto Q_k pertence ao segmento de reta em $\mathbb{R}^n$ cujos extremos são R_k e S_k. Como $\lim R_k = \lim S_k = f'(c)$, concluimos que $\lim Q_k = f'(c)$, o que prova o lema.

Lema 2. *Sejam $f\colon [a, b] \to \mathbb{R}^n$ e $\varphi\colon [a, b] \to \mathbb{R}$ contínuas, e diferenciáveis no intervalo aberto (a, b). Se $|f'(t)| \leq \varphi'(t)$ e $\varphi'(t) > 0$ para todo $t \in (a, b)$ então $|f(b) - f(a)| \leq \varphi(b) - \varphi(a)$.*

Demonstração: Suponhamos inicialmente que f e φ sejam diferenciáveis no intervalo fechado $[a, b]$ e admitamos, por absurdo, que $|f(b) - f(a)| > \varphi(b) - \varphi(a)$. Então existe $A > 1$ tal que $|f(b) - f(a)| > A \cdot [\varphi(b) - \varphi(a)]$. Neste caso, dividindo o intervalo $[a, b]$ ao meio, pelo menos uma das metades, digamos $[a_1, b_1]$, seria tal que $|f(b_1) - f(a_1)| > A \cdot [\varphi(b_1) - \varphi(a_1)]$. Analogamente, pelo menos uma das metades do intervalo $[a_1, b_1]$, digamos $[a_2, b_2]$, seria tal que $|f(b_2) - f(a_2)| > A \cdot [\varphi(b_2) - \varphi(a_2)]$. Prosseguindo, obteremos uma seqüência de intervalos $[a, b] \supset [a_1, b_1] \supset \cdots \supset [a_k, b_k] \supset \ldots$ tais que $b_k - a_k = (b - a)/2^k$ e $|f(b_k) - f(a_k)| >$

$A \cdot [\varphi(b_k) - \varphi(a_k)]$ para todo $k \in \mathbb{N}$. Então $\lim a_k = \lim b_k = c \in [a, b]$ e, pelo Lema 1,

$$|f'(c)| = \lim_{k\to\infty} \frac{|f(b_k) - f(a_k)|}{b_k - a_k} \geq A \cdot \lim_{k\to\infty} \frac{\varphi(b_k) - \varphi(a_k)}{b_k - a_k} = \\ = A \cdot \varphi'(c) > \varphi'(c),$$

uma contradição. O caso geral, em que se supõe que φ' exista apenas no interior aberto (a, b), reduz-se ao anterior pois temos $|f(d) - f(c)| \leq \varphi(d) - \varphi(c)$ para todo $[c, d] \subset (a, b)$. Como f e φ são contínuas nos pontos a e b, podemos passar ao limite, fazendo $d \to b$ e $c \to a$ na desigualdade acima, obtendo o Lema 2.

A Desigualdade do Valor Médio segue-se do Lema 2, tomando $\varphi(t) = M \cdot t$.

Observação: A demonstração que demos para o Teorema do Valor Médio aplica-se para qualquer norma que se considere no espaço $\mathbb{R}^n$. Quando a norma provém de um produto interno, há uma terceira demonstração, mais simples, que é a seguinte: a função $\varphi\colon [a, b] \to \mathbb{R}$, definida por $\varphi(t) = \langle f(t), f(b) - f(a)\rangle$, é contínua, e possui em (a, b) a derivada $\varphi'(t) = \langle f'(t), f(b) - f(a)\rangle$. O Teorema do Valor Médio para funções reais fornece $c \in (a, b)$ tal que $\varphi(b) - \varphi(a) = \varphi'(c)(b - a)$, ou seja:

$$|f(b) - f(a)|^2 = \langle f'(c), f(b) - f(a)\rangle \cdot (b - a).$$

A desigualdade de Cauchy-Schwarz nos dá então:

$$|f(b) - f(a)|^2 \leq |f'(c)|\,|f(b) - f(a)|(b - a) \leq M|f(b) - f(a)|(b - a),$$

e daí $|f(b) - f(a)| \leq M(b - a)$.

Corolário. *Se o caminho $f\colon [a, b] \to \mathbb{R}^n$ é contínuo e possui derivada nula em todos os pontos de (a, b) então f é constante.*

Daí resulta, como de costume, que se dois caminhos f, g têm derivadas iguais em todos os pontos, eles diferem por uma constante.

O corolário acima pode também ser demonstrado tomando-se cada coordenada de f e aplicando o resultado análogo para funções reais.

Fórmula de Taylor com resto de Lagrange. *Seja $f\colon [a, a+h] \to \mathbb{R}^n$ um caminho de classe C^{p-1}, p vezes diferenciável no intervalo aberto $(a, a + h)$. Se $|f^{(p)}(t)| \leq M$ para todo $t \in (a, a + h)$ então*

$$f(a + h) = f(a) + h \cdot f'(a) + \cdots + \frac{h^{p-1}}{(p-1)!} f^{(p-1)}(a) + r_p$$

onde $|r_p| \leq M \cdot h^p/p!$.

Pondo $a + h = b$*, a fórmula acima toma o aspecto*

$$f(b) = f(a) + (b-a)f'(a) + \cdots + \frac{(b-a)^{p-1}}{(p-1)!} f^{(p-1)}(a) + r_p$$

onde $|r_p| \leq M \cdot (b-a)^p/p!$

Demonstração: Definamos o caminho $g \colon [a,b] \to \mathbb{R}^n$ pondo

$$g(t) = f(t) + (b-a)f'(t) + \cdots + \frac{(b-t)^{p-1}}{(p-1)!} f^{(p-1)}(t).$$

Então g é contínuo em $[a,b]$ e diferenciável em (a,b), com $g'(t) = (b-a)^{p-1} \cdot f^{(p)}(t)/(p-1)!$. Logo $|g'(t)| \leq M(b-t)^{p-1}/(p-1)!$. Pondo $\varphi(t) = -M(b-t)^p/p!$, o Lema 2 se aplica, resultando $|g(b)-f(a)| \leq \varphi(b)-\varphi(a)$, ou seja, $|r_p| \leq M(b-a)^p/p!$.

Comentário sobre as três fórmulas de Taylor

Cada uma delas generaliza, para caminhos p vezes diferenciáveis, uma propriedade conhecida de caminhos apenas uma vez diferenciáveis.

A fórmula de Taylor infinitesimal fornece uma aproximação polinomial análoga à aproximação linear que é essencialmente a definição de derivada. Nela, supõe-se a existência das $p-1$ primeiras derivadas de f numa vizinhança de a mas a p-ésima derivada $f^{(p)}$ é suposta existir apenas no ponto a. A fórmula em si é apenas uma definição do "resto"$r_p(h)$. A afirmação importante nela é que $r_p(h)/h^p$ tende para zero juntamente com h, ou seja, que o resto $r_p(h)$ é um infinitésimo de ordem $\geq p$ em relação a h. Tal fórmula, portanto, só diz alguma coisa quando fazemos h tender para zero. Daí o seu nome.

A fórmula de Taylor com resto de Lagrange estende, para caminhos p vezes diferenciáveis, o Teorema do Valor Médio. Neste caso, o acréscimo h é fixo; não se tem a mínima intenção de fazê-lo tender a zero. A fórmula dá uma informação sobre o valor $f(a+h)$ em termos das derivadas $f^{(i)}(a)$, $0 \leq i < p$, e do comportamento de $f^{(p)}(t)$ em todo o intervalo $(a, a+h)$. Daí o nome tradicional de "teorema dos acréscimos finitos". Aqui, a palavra "finito"está em oposição a "infinitésimo", não a "infinito"; ela visa informar que h não vai tender para zero.

Finalmente, a fórmula de Taylor com resto integral é o análogo, para caminhos p vezes diferenciável, do Teorema Fundamental do Cálculo.

Ela tem um caráter "finito", como a fórmula com resto de Lagrange, com a vantagem de exprimir o resto através de uma igualdade, sob a forma explícita de uma integral. Por outro lado, ela impõe à derivada $f^{(p)}$ a condição de ser integrável.

4 Caminhos retificáveis

Definiremos o comprimento do caminho $f\colon [a,b] \to \mathbb{R}^n$ como a distância total percorrida pelo ponto móvel $f(t)$, quando o parâmetro t (que podemos imaginar como o tempo) varia de a até b. Não se trata do comprimento da imagem $f([a,b])$: para ir de $f(a)$ até $f(b)$, o ponto $f(t)$ pode passar pelo mesmo trecho várias vezes. Por exemplo, a imagem do caminho $f\colon [-\sqrt{\pi}, \sqrt{\pi}] \to \mathbb{R}^2$, onde $f(t) = (\cos t^2, \operatorname{sen} t^2)$, é o semicírculo $X = \{(x,y) \in S^1; y \geq 0\}$, cujo comprimento é π. Entretanto $f(t)$ percorre X duas vezes quando t varia de $-\sqrt{\pi}$ até $\sqrt{\pi}$. Portanto o comprimento do caminho f, como veremos adiante, é 2π.

Distinguiremos os caminhos retificáveis (para os quais a medida do percurso é finita) dos não-retificáveis (que percorrem uma distância infinita num tempo finito). Admitiremos caminhos descontínuos, mas a continuidade não assegura retificabilidade: como será visto abaixo, quando $a > 0$, o caminho $f\colon [-a,a] \to \mathbb{R}$, definido por $f(t) = t \operatorname{sen} \dfrac{1}{t}$ se $t \neq 0$ e $f(0) = 0$, não é retificável. Note-se que, par todo $t \in [-a,a]$, o ponto $f(t)$ pertence ao intervalo $[-a,a]$. Entretanto, quando t se aproxima de zero, $f(t)$ oscila infinitas vezes em torno de 0 e, nesse movimento oscilatório, a distância total percorrida é infinita. Observemos ainda que, para ter comprimento infinito, o caminho não precisa passar infinitas vezes pelo mesmo lugar; a espiral $f\colon [0,1] \to \mathbb{R}^2$, onde $f(t) = te^{i/t}$ se $t \neq 0$ e $f(0) = 0$, é um caminho injetivo não-retificável.

Seja $f\colon [a,b] \to \mathbb{R}^n$ um caminho. A cada partição $P = \{t_0, \ldots, t_k\}$ do intervalo $[a,b]$, associaremos o número

$$\ell(f;P) = |f(t_1) - f(t_0)| + |f(t_2) - f(t_1)| + \cdots + |f(t_k) - f(t_{k-1})| =$$
$$= \sum_{i=1}^{k} |f(t_i) - f(t_{i-1})|.$$

Quando não houver perigo de confusão, escreveremos $\ell(P)$, em vez de $\ell(f;P)$. Intuitivamente, $\ell(P)$ é o comprimento da poligonal inscrita no caminho f, com vértices nos pontos $f(t_i)$.

Dizemos que uma partição Q do intervalo $[a,b]$ é *mais fina* do que outra partição P do mesmo intervalo quando $P \subset Q$. Mostraremos abaixo que, refinando-se uma partição, o comprimento da poligonal inscrita correspondente não diminui.

Teorema 1. *Se $P \subset Q$ então $\ell(P) \leq \ell(Q)$.*

Demonstração: Suporemos que $Q = P \cup \{r\}$, com $t_{i-1} \leq r \leq t_i$. (O caso geral reduz-se a este, aplicado várias vezes.) Então $\ell(Q) - \ell(P) = |f(t_i) - f(r)| + |f(r) - f(t_{i-1})| - |f(t_i) - f(t_{i-1})|$. Pela desigualdade triangular, temos $\ell(Q) - \ell(P) \geq 0$.

Dado o caminho $f\colon [a,b] \to \mathbb{R}^n$, façamos P variar entre todas as partições do intervalo $[a,b]$. Se o conjunto dos números $\ell(P)$, assim obtidos, for limitado, diremos que f é um caminho *retificável* e ao número $\ell(f) = \sup\limits_P \ell(P)$, extremo superior dos comprimentos das poligonais inscritas em f, chamaremos o *comprimento* do caminho f.

Assim, o comprimento $\ell(f)$ do caminho retificável $f\colon [a,b] \to \mathbb{R}^n$ é caracterizado por duas propriedades:

1) $\ell(f) \geq \ell(P)$ para toda partição P de $[a,b]$;
2) Dado $\varepsilon > 0$, existe uma partição P tal que $\ell(P) > \ell(f) - \varepsilon$.

Quando $n = 1$, um caminho retificável $f\colon [a,b] \to \mathbb{R}$ chama-se uma função de *variação total limitada* e o comprimento $\ell(f)$ chama-se a *variação total* da função f no intervalo $[a,b]$.

Todo caminho retificável $f\colon [a,b] \to \mathbb{R}^n$ é limitado. Com efeito, para todo $t \in [a,b]$, consideremos a partição $P = \{a,t,b\}$. Então $|f(t) - f(a)| \leq |f(a) - f(t)| + |f(t) - f(b)| = \ell(P) \leq \ell(f)$. Assim, todos os pontos $f(t)$, pertencentes à imagem de f, estão contidos na bola fechada de centro $f(a)$ e raio igual ao comprimento $\ell(f)$.

Fixemos uma partição P_0 do intervalo $[a,b]$. Se, na definição do comprimento de um caminho $f\colon [a,b] \to \mathbb{R}^n$, em vez de todas as partições de $[a,b]$, considerarmos apenas aquelas que são mais finas do que P_0, obteremos o mesmo resultado. Este é o conteúdo do

Lema. $\sup\limits_P \ell(P) = \sup\limits_{Q \supset P_0} \ell(Q)$.

Demonstração: Evidentemente, vale $\geq$ acima. Por outro lado, dada qualquer partição P, considerando $Q = P \cup P_0$ temos $Q \supset P_0$ e $Q \supset P$. Logo, pelo Teorema 1, $\ell(P) \leq \ell(Q)$. Segue-se que $\sup\limits_P \ell(P) \leq \sup\limits_{Q \supset P_0} \ell(Q)$, completando assim a demonstração do lema.

Usaremos o lema para mostrar que se um caminho é formado pela justaposição de dois outros, com um ponto em comum, seu comprimento é igual à soma dos comprimentos desses dois caminhos parciais.

Teorema 2. *Dado $c \in [a,b]$, o caminho $f\colon [a,b] \to \mathbb{R}^n$ é retificável se, e somente se, suas restrições $f_1 = f|[a,c]$ e $f_2 = f|[c,b]$ são retificáveis. No caso afirmativo, tem-se $\ell(f) = \ell(f_1) + \ell(f_2)$.*

Demonstração: Uma partição P do intervalo $[a,b]$ contém o ponto c se, e somente se, $P = P_1 \cup P_2$ onde P_1 é uma partição de $[a,c]$ e P_2 uma partição de $[c,b]$. No caso afirmativo, tem-se $\ell(f;P) = \ell(f_1;P_2) + \ell(f_2;P_2)$. Do lema anterior, segue-se que

$$\ell(f) = \sup_{c\omega`} \ell(f;P) = \sup_{p_1} \ell(f_1;P_1) + \sup_{P_2} \ell(f_2;P_2) = \ell(f_1) + \ell(f_2).$$

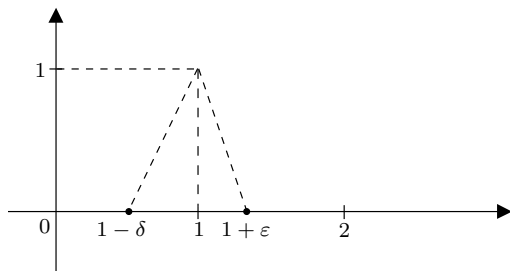
Exemplo 8. Seja $f\colon [0,1] \to \mathbb{R}^n$ o caminho retilíneo $f(t) = (1-t)A + tB$, com $A, B \in \mathbb{R}^n$. Sua imagem é o segmento de reta $[A,B]$. Para toda partição P do intervalo $[0,1]$, temos $\ell(P) = |B-A|$, logo o comprimento do caminho retilíneo f tal que $f(0) = A$ e $f(1) = B$ é igual a $|B-A|$. Em geral, se um caminho $f\colon [a,b] \to R^n$ tem extremidades $A = f(a)$ e $B = f(b)$, então $\ell(f) \geq |B-A|$, pois $P = \{a,b\}$ é uma partição tal que $\ell(P) = |B-A|$. No caso extremo em que $\ell(f) = |B-A|$, podemos assegurar que sua imagem $f([a,b])$ está contida no segmento de reta $[A,B]$, *desde que a norma em $\mathbb{R}^n$ provenha de um produto interno.* Com efeito, se existisse $c \in [a,b]$ tal que $C = f(c) \notin [A,B]$ então $|B-A| < |B-C| + |C-A|$ logo $\ell(f) \geq |B-C| + |C-A| > |B-A|$. Quando f é contínua, a inclusão $f([a,b]) \subset [A,B]$ é necessariamente uma igualdade pois $f([a,b])$, sendo um subconjunto conexo de $[A,B]$, é um segmento de reta. Como contém A e B, tal segmento é igual a $[A,B]$. Note-se porém que, se a norma não provém de um produto interno, podemos ter $|B-A| = |B-C| + |C-A|$ sem que $C \in [A,B]$, o que permite a existência de um caminho f, com $\ell(f) = |f(b) - f(a)|$, sem que a imagem de f seja um segmento de reta. Basta tomar a norma da soma em $\mathbb{R}^2$ e o caminho poligonal f, formado pelo segmento $[0,1]$ do eixo das abcissas mais o segmento $[0,1]$ do eixo das ordenadas. Temos $\ell(f) = 2$. As extremidades de f são os pontos $A = (0,1)$ e $B = (1,0)$, que cumprem $|B-A| = 2$, na norma da soma. Assim $\ell(f) = |B-A|$ embora a imagem de f não seja um segmento de reta.

Evidentemente, $\ell(f) = 0 \Leftrightarrow f$ é um caminho constante.

Ser ou não ser retificável é uma propriedade do caminho f que não depende da norma tomada em $\mathbb{R}^n$, já que duas normas quaisquer são equivalentes.

Com efeito, se $|x| \leq c||x||$ para todo $x \in \mathbb{R}^n$ e $c > 0$ constante então, para toda partição P, temos $\Sigma|f(t_i) - f(t_{i-1})| \leq c \cdot \Sigma||f(t_i) - f(t_{i-1})||$. Assim, se f é retificável segundo a norma $|\ |$ é também retificável segundo $||\ ||$. O comprimento $\ell(f)$ entretanto depende da norma. O segmento de reta que liga os pontos $A = (0, 1)$ e $B = (1, 0)$ no plano, tem comprimento 2 na norma da soma e $\sqrt{2}$ na norma euclidiana.

Exemplos: **9**. Se f é um caminho poligonal, seu comprimento é, de acordo com o Teorema 2, igual à soma dos comprimentos dos segmentos que o compõem. Em particular, para toda partição P, o número $\ell(f; P) = \Sigma|f(t_i) - f(t_{i-1})|$ é, realmente, o comprimento da poligonal inscrita em f, cujos vértices são os pontos $f(t_i)$.

10. O caminho $f\colon [0, 2] \to \mathbb{R}^2$, definido por $f(t) = (t, 0)$ para $t \neq 1$ e $f(1) = (1, t)$, é descontínuo, porém retificável. Tomando a norma euclidiana em $\mathbb{R}^2$, temos $\ell(f) = 4$. Para verificar isto, basta considerar partições P que contém o ponto $t_i = 1$.

Se $1 - \delta$ é o último ponto antes de 1 e $1 + \varepsilon$ o primeiro depois de 1 na partição P, temos $\ell(P) = 1 - \delta + \sqrt{1 + \delta^2} + \sqrt{1 + \varepsilon^2} + 1 - \varepsilon$. Isto mostra que $\ell(P) \leq 4$ e que, tomando δ, ε suficientemente pequenos, podemos tornar $\ell(P)$ tão próximo de 4 quanto desejamos. Logo $\ell(f) = 4$.

Teorema 3. *O caminho $f\colon [a, b] \to \mathbb{R}^n$ é retificável se, e somente se, cada uma das suas funções coordenadas $f_i\colon [a, b] \to \mathbb{R}$ é retificável (isto é, tem variação total limitada).*

Demonstração: Como a retificabilidade de f independe da norma, tomemos em $\mathbb{R}^n$ a norma da soma. Então, para toda partição P do intervalo $[a, b]$, temos $\ell(f; P) = \sum\limits_{1 \leq i \leq n} \ell(f_i; P)$. Portanto, quando P varia,

o conjunto dos números $\ell(f;P)$ é limitado se, e somente se, é limitado cada um dos conjuntos de números $\ell(f_1;P),\ldots,\ell(f_n;P)$.

Corolário. *Se cada função coordenada do caminho f é monótona então f é retificável.*

Exemplo 11. O caminho contínuo $f\colon [0,1]\to\mathbb{R}$, definido por $f(0)=0$ e $f(t)=t\ \mathrm{sen}(\pi/2t)$ para $t\neq 0$, não é retificável. Com efeito, para t sucessivamente igual a $1,1/2,1/3,\ldots,1/4m$, $f(t)$ assume a seqüência de valores $1,0,-1/3,0,1/5,0,\ldots,-1/(4m-1),0$. Portanto, tomando $k=4m-1$ a partição $P=\{(0,1/(k+1),k,\ldots,1/3,1/2,1\}$ fornece uma soma

$$\ell(P)=\frac{1}{k}+\frac{1}{k}+\cdots+\frac{1}{5}+\frac{1}{5}+\frac{1}{3}+\frac{1}{3}+1,$$

donde

$$\ell(P)\geq\frac{1}{k+1}+\frac{1}{k}+\cdots+\frac{1}{6}+\frac{1}{5}+\frac{1}{4}+\frac{1}{3}+\frac{1}{2}\cdot$$

Como a série harmônica diverge, vemos que existem partições P que tornam $\ell(P)$ tão grande quanto desejemos, portanto f não é retificável. Resulta do Teorema 3 que tampouco é retificável o caminho plano $f\colon[0,1]\to\mathbb{R}^2$, onde $f(t)=(t,t\ \mathrm{sen}(1/t))$ se $t\neq 0$ e $f(0)=(0,0)$. (Note que f é injetivo.) Isto se exprime dizendo que o gráfico da função contínua $t\mapsto t\cdot\mathrm{sen}(1/t)$ tem comprimento infinito se $0\leq t\leq 1$. Também tem comprimento infinito o caminho espiralado $f\colon[0,1]\to\mathbb{R}^2$, dado por $f(t)=(t\cos 1/t,\ t\ \mathrm{sen}\, 1/t)=t\cdot e^{i/t}$. Neste caso, quando $t\to 0$, o ponto $f(t)$ tende para a origem $(0,0)$ dando voltas em torno dela, de modo a percorrer uma rota infinitamente longa antes de atingir a meta $f(0)$.

Já vimos no Exemplo 10 que um caminho descontínuo pode ser retificável. Mas a descontinuidade de um caminho retificável $f\colon[a,b]\to\mathbb{R}^n$ num ponto $c\in[a,b]$ não pode ser arbitrária. Quando $t\to c$, não podem existir dois valores de aderência distintos A e B para o caminho f. Com efeito, para oscilar indefinidamente entre A e B, o ponto variável $f(t)$ teria que percorrer infinitas vezes uma distância aproximadamente igual a $|B-A|$ e então f não seria retificável. De modo preciso, temos o

Teorema 4. *Seja $f\colon[a,b)\to\mathbb{R}^n$ um caminho tal que, para cada $c\in[a,b)$, a restrição $f|[a,c]$ é retificável. Se existir $K>0$ com $\ell(f|[a,c])\leq K$ para todo $c\in[a,b]$ então existe $\lim\limits_{t\to b^-}f(t)$.*

Analogamente, dado $f\colon(a,b]\to\mathbb{R}^n$ tal que $f|[c,b]$ é retificável para todo $c\in(a,b]$, com $\ell(f|[c,b])\leq K$ seja qual for $c\in(a,b]$, então existe $\lim\limits_{t\to a^+}f(t)$.

Demonstração: Basta demonstrar a primeira parte. Seja então $t_1 < t_2 < \cdots < t_k < \ldots$ uma seqüência crescente em $[a,b)$, com $\lim t_k = b$. Para todo $k \in \mathbb{N}$, temos $\sum_{i=2}^{k} |f(t_i) - f(t_{i-1})| \leq K$ pois $\{a, t_1, \ldots, t_k\}$ é uma partição de $[a,c]$, com $c = t_k$. A série de números reais $\sum_{i=2}^{\infty} |f(t_i) - f(t_{i-1})|$ é portanto convergente. Pelo critério de Cauchy, a série de vetores $\sum_{i\geq 2} [f(t_i) - f(t_{i-1})]$ também é convergene. Como a reduzida de ordem $k-1$ desta série é igual a $f(t_k) - f(t_1)$, segue-se que existe $\lim f(t_k)$. Como a seqüência crescente (t_k) é arbitrária, concluimos que existe $\lim_{t\to b^-} f(t)$.

Corolário. *Seja $f\colon [a,b] \to \mathbb{R}^n$ retificável. Para todo $c \in [a,b]$, existem os limites laterais $\lim_{t\to c^-} f(t)$ (se $c \neq a$) e $\lim_{t\to c^+} f(t)$ (se $c \neq b$).*

Os caminhos com a propriedade expressa pelo corolário acima chama-se *regulados*. O conjunto dos pontos de descontinuidade de um caminho regulado é enumerável. (V. vol. 1, p. 233, Teorema 11.)

Usaremos as notações $f(c_+) = \lim_{t\to c^+} f(t)$ e $f(c^-) = \lim_{t\to c^-} f(t)$.

Motivados pela integral, podemos indagar se o comprimento de um caminho f pode ser definido como $\ell(f) = \lim_{|P|\to 0} \ell(f;P)$, esta igualdade significando que, para todo $\varepsilon > 0$ dado, pode-se obter $\delta > 0$ tal que $|P| < \delta \Rightarrow |\ell(f;P) - \ell(f)| < \varepsilon$.

A caracterização do comprimento como limite, entretanto, só é válida para uma classe especial de caminhos, que introduziremos agora.

Diremos que um caminho $f\colon [a,b] \to \mathbb{R}^n$ é *bem regulado* quando ele for regulado e, para todo $c \in (a,b)$,

$$|f(c^+) - f(c^-)| = |f(c^+) - f(c)| + |f(c) - f(c^-)|.$$

Quando a norma provém de um produto interno, a igualdade acima significa que o ponto $f(c)$ pertence ao segmento de reta cujos extremos são $f(c^-)$ e $f(c^+)$. Para uma norma arbitrária, podemos apenas dizer que se $f(c) \in [f(c^-), f(c^+)]$ para todo $c \in (a,b)$ então f é bem regulado.

Todo caminho contínuo é bem regulado. Todo caminho regulado lateralmente contínuo (isto é $f(c) = f(c^+)$ ou $f(c) = f(c^-)$ para todo $c \in (a,b)$) é bem regulado. Um caminho é bem regulado se, e somente

se, para todo $c \in (a, b)$ e $s < c < t$, vale:

$$\lim_{\substack{s \to c^- \\ t \to c^+}} [|f(t) - f(c)| + |f(c) - f(s)| - |f(t) - f(s)|] = 0.$$

Exemplo 12. O caminho retificável do Exemplo 10 não é bem regulado. Para ele, não existe $\lim\limits_{|P| \to 0} \ell(P)$, pois se a partição P não contém o ponto $t = 1$, vale sempre $\ell(P) = 2$, enquanto que, para partições Q que contêm esse ponto, temos $\lim\limits_{|Q| \to 0} \ell(Q) = 4$.

Teorema 5. *As seguintes afirmações a respeito de um caminho $f\colon [a, b] \to \mathbb{R}^n$ são equivalentes:*

(1) *f é bem regulado e retificável, com $\ell(f) = L$;*

(2) *existe $\lim\limits_{|P| \to 0} \ell(f; P) = L$.*

Demonstração: (1) $\Rightarrow$ (2). Com efeito, dado $\varepsilon > 0$, seja $P_0 = \{t_0, t_1, \ldots, t_k\}$ uma partição tal que $L - \varepsilon/2 < \ell(P_0) \leq L$. Tomemos $\delta > 0$ menor do que o comprimento do menor dos intervalos de P_0 e tal que $t_i - \delta < s < t_i < t < t_i + \delta \Rightarrow ||f(t) - f(t_i)| + |f(t_i) - f(s)| - |f(t) - f(s)|| < \varepsilon/2k$ para todo $i = 1, 2, \ldots, k - 1$. Afirmamos que $|P| < \delta \Rightarrow |\ell(P) - L| < \varepsilon$. Com efeito, dada a partição P com norma menor do que δ, temos $L - \varepsilon/2 < \ell(P \cup P_0) \leq L$ e $0 \leq \ell(P \cup P_0) - \ell(P) =$ soma de, no máximo k, termos da forma $|f(t) - f(t_i)| + |f(t_i) - f(s)| - |f(t) - f(s)|$, onde $[s, t]$ é um intervalo de P que contém algum ponto t_i em seu interior. (Os demais intervalos de P são também de $P \cup P_0$, portanto desaparecem na diferença $\ell(P \cup P_0) - \ell(P)$.) Segue-se que $|P| < \delta$ implica $0 \leq \ell(P \cup P_0) - \ell(P) < \varepsilon/2$ o que, por sua vez, acarreta $L - \varepsilon < \ell(P) \leq L$, como queríamos demonstrar.

(2) $\Rightarrow$ (1). Dado arbitrariamente $\varepsilon > 0$ existe, pela hipótese, $\delta > 0$ tal que $|P| < \delta \Rightarrow L - \varepsilon < \ell(P) < L + \varepsilon$. Fixemos uma partição P_0, com norma menor do que δ. Como $P \supset P_0 \Rightarrow |P| \leq |P_0| < \delta$, temos então $L - \varepsilon < \ell(P) < L + \varepsilon$ sempre que $P \supset P_0$. Resulta assim do Lema anterior ao Teorema 2 que f é retificável e que $L - \varepsilon \leq \ell(f) \leq L + \varepsilon$. Sendo $\varepsilon > 0$ arbitrário, concluimos que $\ell(f) = L$. Em seguida, mostraremos que a existência do limite $L = \lim\limits_{|P| \to 0} \ell(P)$ implica que f é bem regulado. Com efeito, dado um ponto arbitrário $c \in (a, b)$, tomemos

um seqüência de partições Q_k tais que $\lim |Q_k| = 0$ e $c \notin Q_k$ para todo k. Ponhamos $P_k = Q_k \cup \{c\}$. Então $\lim \ell(Q_k) = \lim \ell(P_k) = L$ e

$$(*) \qquad \ell(P_k) - \ell(Q_k) = |f(t_k) - f(c)| + |f(c) - f(s_k)| - \\ - |f(t_k) - f(s_k)|,$$

onde $[s_k, t_k]$ é o intervalo da partição Q_k que contém c em seu interior. Evidentemente, $\lim f(t_k) = f(c_+)$ e $\lim f(s_k) = f(c^-)$. Fazendo $k \to \infty$ em (*) obtemos então:

$$0 = \lim[\ell(P_k) - \ell(Q_k)] = |f(c^+) - f(c)| + \\ + |f(c) - f(c^-)| - |f(c^+) - f(c^-)|,$$

logo f é bem regulado.

Corolário. *Um caminho contínuo $f\colon [a,b] \to \mathbb{R}^n$ tem comprimento L se, e somente se, $\lim\limits_{|P|\to 0} \ell(P) = L$.*

Se o caminho $f\colon [a,b] \to \mathbb{R}^n$ é Lipschitziano, com $|f(s) - f(t)| \leq k|s-t|$ para $s,t \in [a,b]$ quaisquer, então, para cada partição P do intervalo $[a,b]$, tem-se $\ell(P) \leq k\cdot(b-a)$. Segue-se que f é retificável e seu comprimento $\ell(f)$ não excede $k\cdot(b-a)$. Um caso particular importante dá-se quando $f \in C^1$. Neste caso, o teorema abaixo fornece uma expressão para o comprimento de f. (Veja também Exercício 4.12.)

Teorema 6. *Todo caminho $f\colon [a,b] \to \mathbb{R}^n$ de classe C^1 é retificável, com $\ell(f) = \int_a^b |f'(t)|\,dt$.*

Demonstração: Mostraremos que $\lim\limits_{|P|\to 0} \ell(P) = \int_a^b |f'(t)|\,dt$. Seja, pois, dado $\varepsilon > 0$. Pela definição de integral, se pontilharmos a partição $P = \{t_0, t_1, \ldots, t_k\}$ tomando sempre $\xi_i = t_{i-1} \in [t_{i-1}, t_i]$, veremos que existe $\delta_1 > 0$ tal que $|P| < \delta_1$ implica

$$\left| \int_a^b |f'(t)|\,dt - \sum_{i=1}^{k} |f'(t_{i-1})|(t_i - t_{i-1}) \right| < \varepsilon/2.$$

Além disso, pela diferenciabilidade uniforme de f, existe $\delta_2 > 0$ tal que $|P| < \delta_2$ implica

$$f(t_i) - f(t_{i-1}) = [f'(t_{i-1}) + \rho_i](t_i - t_{i-1}), \text{ com } |\rho_i| < \frac{\varepsilon}{2(b-a)}\cdot$$

Logo, $|P| < \delta_2 \Rightarrow |\ell(P) - \Sigma|f'(t_{i-1})|(t_i t_{i-1})| < \varepsilon/2$. Seja $\delta = \min\{\delta_1, \delta_2\}$. Então $|P| < \delta \Rightarrow |\ell(P) - \int_a^b |f'(t)|dt| < \varepsilon$, o que demonstra o teorema.

Exemplo 13. Seja $f\colon [0, 2\pi] \to \mathbb{R}^2$, $f(t) = (\cos t, \operatorname{sen} t)$. Então o comprimento de f é igual a $\int_0^{2\pi} |f'(t)|\, dt = \int_0^{2\pi} 1 \cdot dt = 2\pi$. Analogamente, o caminho $g\colon [-\sqrt{\pi}, \sqrt{\pi}] \to \mathbb{R}^2$, $g(t) = f(t^2) = (\cos t^2, \operatorname{sen} t^2)$, tem comprimento igual a

$$\int_{-\sqrt{\pi}}^{\sqrt{\pi}} |g'(t)|\, dt = \int_{-\sqrt{\pi}}^{\sqrt{\pi}} |2t|\, dt = 2\int_0^{\sqrt{\pi}} 2t\, dt = 2t^2\Big]_0^{\sqrt{\pi}} = 2\pi.$$

5 O comprimento de arco como parâmetro

Seja $g\colon [c, d] \to \mathbb{R}^n$ um caminho. Uma *reparametrização* de g é um caminho $g \circ \varphi\colon [a, b] \to \mathbb{R}^n$, onde $\varphi\colon [a, b] \to [c, d]$ é uma função monótona sobrejetiva. (Pelo Teor. 10, pág. 232, vol. 1, φ é então contínua.) Quando φ é não-decrescente, tem-se $\varphi(a) = c$, $\varphi(b) = d$ mas se φ é não-crescente, então $\varphi(a) = d$, $\varphi(b) = c$. Não estamos supondo que φ seja injetiva mas, em virtude de sua monotonicidade, se $\varphi(s) = \varphi(t)$ com $s < t$ então φ é constante no intervalo $[s, t]$.

Observação: Se a reparametrização $f = g \circ \varphi\colon [a, b] \to \mathbb{R}^n$ é contínua, então o caminho $g\colon [c, d] \to \mathbb{R}^n$ é contínuo. Isto decorre de (12.6), no §12 do Capítulo 1.

Teorema 7. *A reparametrização $g \circ \varphi\colon [a, b] \to \mathbb{R}^n$ é retificável se, e somente se, o caminho $g\colon [c, d] \to \mathbb{R}^n$ é retificável. No caso afirmativo, tem-se $\ell(g \circ \varphi) = \ell(g)$.*

Demonstração: Suponhamos g retificável. Dada uma partição $P = \{s_0, \ldots, s_k\}$ de $[a, b]$, se existirem s_{i-1}, $s_i \in P$ com $\varphi(s_{i-1}) = \varphi(s_i)$ então, omitindo o ponto s_i, obtemos uma partição $P^* = P - \{s_i\}$, tal que $\ell(g \circ \varphi; P^*) = \ell(g \circ \varphi; P)$. Portanto, ao computarmos o comprimento de $g \circ \varphi$, basta considerar as partições P de $[a, b]$ tais que $\varphi|P$ seja injetiva. Então $Q = \varphi(P)$ é uma partição de $[c, d]$, tal que $\ell(g \circ \varphi; P) = \Sigma|g\varphi(s_i) - g\varphi(s_{i-1})| = \ell(g; Q) \leq \ell(g)$ para toda P, logo $g \circ \varphi$ é retificável e $\ell(g \circ \varphi) \leq \ell(g)$. Reciprocamente, seja $g \circ \varphi$ retificável. Para toda partição $Q = \{t_0, \ldots, t_k\}$ de $[c, d]$, como φ é sobrejetiva, existe $P = \{s_0, \ldots, s_k\} \subset [a, b]$ tal que $\varphi(s_i) = t_i$ para $i = 0, \ldots, k$. Se φ for não-decrescente (como podemos obviamente tomar $s_0 = a$ e $s_k = b$), P é uma partição de $[a, b]$ e, evidentemente, $\ell(g \circ \varphi; P) = \ell(g; Q)$. Se φ

for não-decrescente, numeraremos os s_i de modo que $\varphi(s_{k-i}) = t_i$ e valerá a mesma conclusão. Assim, para cada partição Q de $[c,d]$, temos $\ell(g;Q) = \ell(g\circ\varphi;P) \leq \ell(g\circ\varphi)$, donde g é retificável e $\ell(g) \leq \ell(g\circ\varphi)$. Concluimos que $\ell(t) = \ell(g\circ\varphi)$.

Um caminho retificável $f\colon [a,b] \to \mathbb{R}^n$ diz-se *cadenciado* (ou parametrizado por comprimento de arco) quando, para todo $t \in [a,b]$, tem-se $\ell(f|[a,t]) = t-a$. Noutros termos: para ir de $f(a)$ a $f(t)$, ao longo do caminho f, percorre-se uma distância igual a $t-a$. Neste caso, se $s<t$ pertencem a $[a,b]$, então $\ell(f|[s,t]) = t-s$.

Teorema 8. *Um caminho $f\colon [a,b] \to \mathbb{R}^n$, de classe C^1, é cadenciado se, e somente se, $|f'(t)| = 1$ para todo $t \in [a,b]$.*

Demonstração: Se f é cadenciado, então $\int_a^t |f'(s)|\,ds = t-a$ para todo $t \in [a,b]$. Então $|f'(t)| = \dfrac{d}{dt}\int_a^t |f'(s)|\,ds = \dfrac{d}{dt}(t-a) = 1$. Reciprocamente, se $|f'(t)| = 1$ para tdo $t \in [a,b]$, então $\ell(f|[a,t]) = \int_a^t |f'(s)|\,ds = \int_a^t 1\cdot ds = t-a$.

Exemplo 14. O caminho $f\colon [a,b] \to \mathbb{R}^2$ $f(t) = (\cos t, \operatorname{sen} t)$, é cadenciado pois $|f'(t)| = |(-\operatorname{sen} t, \cos t)| = 1$.

Teorema 9. *Todo caminho contínuo retificável $f\colon [a,b] \to \mathbb{R}^n$ é a reparametrização de um caminho cadenciado $g\colon [0,L] \to \mathbb{R}^n$, $L = \ell(f)$, o qual é necessariamente contínuo.*

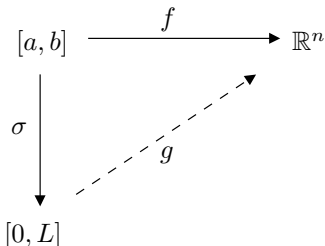
A demonstração se baseia no

Lema. *Se $f\colon [a,b] \to \mathbb{R}^n$ é contínuo e retificável, então a função $\sigma\colon [a,b] \to [0,L]$, $L = \ell(f)$, definida por $\sigma(t) = \ell(f|[a,t])$, é contínua. Como $\sigma(b) = L$, σ é sobrejetiva.*

Demonstração: Basta provar que σ é contínua no ponto a. O caso geral é análogo. Sendo σ monótona, existe $A = \lim\limits_{t\to a^+} \sigma(t) = \inf\limits_{t\geq a} \sigma(t)$. Suponhamos, por absurdo, que fosse $A > 0$. Então poderíamos achar $c \in (a,b)$ tal que $A \leq \sigma(c) < 4A/3$ e daí $A \leq \sigma(t) \leq \sigma(c) < 4A/3$ para todo $t \in [a,c]$. Logo, $\ell(f|[t,c]) = \sigma(c) - \sigma(t) < A/3$. Sendo f contínua, podemos supor c tão próximo de a que $t \in [a,c] \Rightarrow |f(t) - f(a)| < A/3$. Então, para toda partição P do intervalo $[a,c]$ teríamos

$$\ell(P) = |f(t_1) - f(a)| + \sum_{i>2} |f(t_i) - f(t_{i-1})| <$$
$$< \frac{A}{3} + \ell(f|[t_1,c]) < \frac{A}{3} + \frac{A}{3}\,.$$

Logo $\sigma(c) = \ell(f|[a,c]) \leq \dfrac{2A}{3} < A$, uma contradição.

Demonstração do Teorema 9: Consideremos o diagrama abaixo

Sabemos que σ é monótona e sobrejetiva. Trataremos de obter um caminho $g\colon [0,L] \to \mathbb{R}^n$ que torne o diagrama comutativo. Ora, dados $s < t$ em $[a,b]$, temos $\sigma(t) = \sigma(s) + \ell(f|[s,t])$. Portanto $\sigma(s) = \sigma(t) \Rightarrow f$ constante no intervalo $[s,t] \Rightarrow f(s) = f(t)$. Definamos um caminho $g\colon [0,L] \to \mathbb{R}^n$ do seguinte modo: dado $u \in [0,L]$, existe (pelo menos um) $t \in [a,b]$ tal que $\sigma(t) = u$. Pomos $g(u) = f(t)$. A escolha do t não afeta o valor de g, em virtude da observação que acabamos de fazer. Logo, g é bem definido e $f = g \circ \sigma$. A continuidade de g resulta da compacidade de $[a,b]$, já que $f = g \circ \sigma$ e σ são contínuos. [Veja Cap. 1, §12, Observação (12.6)]. Para provar que g é cadenciado tomemos um ponto $s \in [0,L]$ qualquer. Existe $t \in [a,b]$ tal que $s = \sigma(t)$. O Teorema 7 nos permite escrever:

$$\ell(g|[0,s]) = \ell((g \circ \sigma) \mid [a,t]) = \ell(f|[a,t]) = \sigma(t) = s.$$

Isto termina a demonstração.

Corolário. *Um caminho é retificável e, e somente se, é a reparametrização de um caminho Lipschitziano.*

Com efeito, todo caminho Lipschitziano $f\colon [a,b] \to \mathbb{R}$, com $|f(s) - f(t)| \leq k|s-t|$, cumpre $\ell(P) \leq k(b-a)$ para toda partição P, logo é retificável, com $\ell(f) \leq k(b-a)$. Portanto qualquer reparametrização de f é também retificável, pelo Teorema 7. Reciprocamente, se f é contínuo e retificável, então $f = g \circ \sigma$, onde g é cadenciado. Então $|g(t) - g(s)| \leq \ell(g|[s,t]) = |t-s|$, logo g é Lipschitziano (com constante de Lipschitz $k = 1$).

Observações: 1) Seja $f\colon [a,b] \to \mathbb{R}^n$ um caminho contínuo retificável. O caminho cadenciado $g\colon [0,L] \to \mathbb{R}^n$, do qual f é uma reparametrização, é único, porque é única também a função $\sigma\colon [a,b] \to [0,L]$,

monótona e sobrejetiva, tal que $g \circ \sigma = f$. Com efeito, destas propriedades segue-se que para todo $t \in [a, b]$, temos $\sigma(t) = \ell(g|[0, \sigma(t)]) = \ell((g \circ \sigma) \mid [a, t]) = \ell(f|[a, t])$, o que nos leva de volta à definição de σ dada na demonstração do teorema.

2) Um caminho pode ser retificável sem ser Lipschitziano, como por exemplo $f\colon [0, 1] \to \mathbb{R}^2$, $f(t) = (t, \sqrt{t})$. (A retificabilidade de f decorre de serem monótonas suas funções coordenadas.)

Um caminho diferenciável $f\colon I \to \mathbb{R}^n$ diz-se *regular* quando $f'(t) \neq 0$ para todo $t \in I$.

Uma função $f\colon I \to J$ chama-se um *difeomorfismo* entre os intervalos I, J quando é regular e sobrejetiva. Neste caso particular (valores reais), $f'(t) \neq 0$ para todo t obriga (pelo Teorema de Darboux) que o sinal de $f'(t)$ seja constante: ou temos $f'(t) > 0$ para todo t, e f é então monótona crescente ou $f'(t) < 0$ para todo t, com f decrescente. Em particular, *para* $n = 1$, um caminho regular $f\colon I \to \mathbb{R}^n$ é necessariamente injetivo, o que não é verdade quando $n > 1$: por exemplo, $f\colon [a, b] \to \mathbb{R}^2$, $f(t) = (\cos t, \operatorname{sen} t)$, se $b - a > 2\pi$ não é injetivo. Também a ciclóide $f\colon [a, b] \to \mathbb{R}^2$, $f(t) = (t - \operatorname{sen} t, 1 - \cos t)$, é um caminho regular não-injetivo quando $b - a > 2\pi$.

Pelo Teorema da Função Inversa (Vol. 1, pág. 274, Corol. 6), a inversa de um difeomorfismo $f\colon I \to J$ é também um difeomorfismo. Se f for de classe C^k então $f^{-1} \in C^k$.

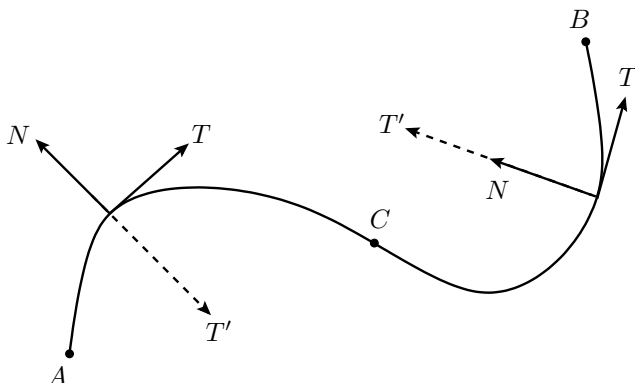
Teorema 10. *Sejam $f\colon [a, b] \to \mathbb{R}^n$ um caminho regular de classe C^k $(k \geq 1)$, $L = \ell(f)$ e $g\colon [0, L] \to \mathbb{R}^n$ o caminho cadenciado, do qual $f = g \circ \sigma$ é uma reparametrização. Então $g \in C^k$ e $\sigma\colon [a, b] \to [0, L]$ é um difeomorfismo de classe C^k. Em particular, $g = f \circ \sigma^{-1}$ também é uma reparametrização de f.*

Demonstração: Temos $\sigma(t) = \ell(f|[a, t]) = \int_a^t |f'(s)|\, ds$, logo $\sigma'(t) = |f'(t)| > 0$ para todo $t \in [a, b]$. Portanto $\sigma\colon [a, b] \to [0, L]$ é um difeomorfismo de classe C^k. Seu inverso, $\sigma^{-1}\colon [0, L] \to [a, b]$, também é de classe C^k, donde $g = f \circ \sigma^{-1} \in C^k$.

6 Curvatura e torção

Seguindo a tradição, indicaremos o parâmetro de um caminho cadenciado com a letra s.

Seja $f\colon [a,b] \to \mathbb{R}^2$ um caminho cadenciado de classe C^2 no plano onde consideramos a *norma euclidiana*. Para cada $s \in [a,b]$, o vetor velocidade $f'(s)$ tem comprimento 1. Escreve-se $f'(s) = T(s)$ e diz-se que $T(s)$ é o vetor *unitário tangente* ao caminho f no ponto $f(s)$.

Como $T(s)$ tem comprimento constante, o vetor $f''(s) = T'(s)$ é perpendicular a $T(s)$. Seja $N(s)$ o vetor unitário obtido de $T(s)$ por meio de uma rotação positiva de 90 graus, isto é, se $T(s) = (\alpha, \beta)$ então $N(s) = (-\beta, \alpha)$. Chama-se $N(s)$ o *vetor unitário normal* ao caminho f no ponto $f(s)$. Existe um único número real $k(s)$ tal que $T'(s) = k(s) \cdot N(s)$. O número $k(s)$ chama-se a *curvatura* do caminho f no ponto $f(s)$. O valor absoluto da curvatura $k(s)$ mede a rapidez com que f muda de direção; seu sinal indica a "concavidade" ou "convexidade" do caminho.

Na figura anterior, esboçamos a imagem de um caminho cadenciado $f\colon [a,b] \to \mathbb{R}^2$, com $f(a) = A$, $f(b) = B$ e $f(c) = C$, $a < c < b$. No trecho AC, temos $k < 0$, enquanto que $k > 0$ no trecho CB. No ponto C, temos $k = 0$; isto é o que se chama um *ponto de inflexão*.

Exemplo 15. Para determinar a curvatura do círculo de centro (a,b) e raio r, usaremos a parametrização $f(s) = (a+r\cdot\cos(s/r), b+r\cdot\operatorname{sen}(s/r))$, a fim de termos $|f'(s)| = 1$. Então $f'(s) = (-\operatorname{sen}(s/r), \cos(s/r))$ e $f''(s) = \frac{1}{r}(-\cos(s/r), -\operatorname{sen}(s/r)) = \frac{1}{r}N(s)$. Portanto, um círculo de raio r tem curvatura constante, igual a $1/r$.

Como o vetor velocidade $T = T(s)$ de um caminho cadenciado $f\colon [a,b] \to \mathbb{R}^2$, de classe C^2, tem comprimento constante, igual a 1, o que varia nele é a direção, a qual pode ser medida pelo ângulo que T faz com uma reta fixa, digamos o eixo das abcissas. É natural esperar

que a curvatura de f seja a derivada desse ângulo. Isto será demonstrado agora.

Para cada $s \in [a, b]$ existe $\theta(s) \in \mathbb{R}$ tal que $f'(s) = T(s) = (\cos\theta(s), \operatorname{sen}\theta(s))$. O número $\theta(s)$ expressa, em radianos, uma das determinações do ângulo que o vetor $T(s)$ forma com o eixo orientado das abcissas. Tal número não é único: acha-se definido a menos de um múltiplo inteiro de 2π.

De acordo com o Teorema 11, que demonstraremos no parágrafo seguinte, é possível escolher, para cada $s \in [a, b]$, o número $\theta(s)$ de tal modo que $T(s) = (\cos\theta(s), \operatorname{sen}\theta(s))$ e a função $\theta\colon [a, b] \to \mathbb{R}$, assim definida, seja de classe C^1.

Então $T'(s) = (-\theta'(s) \cdot \operatorname{sen}\theta(s),\ \theta'(s) \cdot \cos\theta(s)) = \theta'(s) \cdot N(s)$. Portanto $k(s) = \theta(s)$, isto é, a curvatura de um caminho cadenciado no plano é a derivada do ângulo que seu vetor velocidade faz com o eixo orientado das abcissas.

Para caminhos em $\mathbb{R}^3$, a curvatura perde o sinal e surge um novo invariante diferencial: a torção.

Seja, pois, $f\colon [a, b] \to \mathbb{R}^3$ um caminho cadenciado de classe C^3 no espaço $\mathbb{R}^3$.

O vetor unitário tangente $T(s) = f'(s)$ é definido como antes mas a curvatura só se define nos pontos onde $T'(s) \neq 0$. Fazendo esta hipótese, pomos $N(s) = T'(s)/|T'(s)| =$ *vetor unitário normal* no ponto $f(s)$ e $k(s) = |T'(s)| =$ *curvatura*. Note que $k(s) > 0$, por definição.

O plano gerado pelos vetores unitários ortogonais $T = T(s)$ e $N = N(s)$ é afetuosamente denominado o *plano osculador* do caminho f no ponto $f(s)$. O *vetor binormal* $B = B(s)$ é definido pela condição de que a lista ordenada $F = (T, N, B)$ seja uma base ortonormal positiva de $\mathbb{R}^3$. Assim, $|B| = 1$, B é perpendicular ao plano osculador e a matriz 3×3 cujas colunas são as coordenadas de T, N e B, nesta ordem, tem determinante positivo, igual a 1. A base $F = F(s) = (T(s), N(s), B(s))$ chama-se o *triedro de Frenet* do caminho f no ponto $f(s)$.

Em seguida, apresentaremos as equações de Frenet, que desempenham um papel central na Geometria Diferencial das Curvas. Elas exprimem as derivadas T', N' e B' como combinações lineares de T, N e B.

Já sabemos que $T' = k \cdot N$. Além disso, como N tem comprimento constante, $N' \perp N$, donde $N' = x \cdot T + w \cdot B$. Para determinar os coeficientes x, w, derivamos a igualdade $\langle N, T\rangle = 0$, obtendo $\langle N', T\rangle +$

$\langle N, T'\rangle = 0$. Logo $x = \langle N', T\rangle = -\langle N, T'\rangle = -k$.

O coeficiente $w = w(s)$ é chamado a *torção* do caminho f no ponto $f(s)$; temos $w = \langle N', B\rangle = -\langle N, B'\rangle$. Logo, se escrevermos $B' = x \cdot T + y \cdot N$, obteremos $x = \langle B', T\rangle = -\langle B, T'\rangle = 0$ e $y = \langle B', N\rangle = -w$. Podemos então escrever as *equações de Frenet*:

$T' =$		$k \cdot N$	
$N' =$	$-k \cdot T$		$w \cdot B$
$B' =$		$-w \cdot N$	

Pondo $F' = (T', N', B')$, temos $F' = A \cdot F$, onde

$$A = \begin{pmatrix} 0 & k & 0 \\ -k & 0 & w \\ 0 & -w & 0 \end{pmatrix}$$

é uma matriz anti-simétrica.

A torção w mede o desvio do caminho f em relação ao plano osculador. De fato, como veremos agora, um caminho cadenciado f, cuja torção w é identicamente nula, é um caminho plano. Basta observar que $w = 0 \Rightarrow B' = -w \cdot N = 0 \Rightarrow B$ constante $\Rightarrow \dfrac{d}{ds}\langle f(s) - f(a), B\rangle = \langle T, B\rangle = 0 \Rightarrow \langle f(s) - f(a), B\rangle =$ constante. Ora, para $s = a$, temos $\langle f(a) - f(a), B\rangle = 0$. Logo $\langle f(s) - f(a), B\rangle = 0$ para todo s. Assim, os pontos $f(s)$ do caminho f acham-se situados no plano perpendicular ao vetor B, isto é, no plano osculador.

7 A função-ângulo

Seja $z\colon [a, b] \to \mathbb{R}^2$ um caminho tal que $|z(t)| = 1$ para todo $t \in [a, b]$. Então $z(t) \in S^1$ para todo t, logo podemos escrever $z\colon [a, b] \to S^1$. Uma *função-ângulo* para o caminho z é uma função $\theta\colon [a, b] \to \mathbb{R}$ tal que $z(t) = (\cos\theta(t), \operatorname{sen}\theta(t))$ para todo $t \in [a, b]$. Se indicarmos com $\xi\colon \mathbb{R} \to S^1$ a "aplicação exponencial", $\xi(t) = (\cos t, \operatorname{sen} t) = e^{it}$, vemos que θ é uma função-ângulo para o caminho z se, e somente se, $z = \xi \circ \theta$.

Teorema 11. *Todo caminho $z\colon [a, b] \to S^1$ de classe C^r $(r \geq 1)$ possui uma função-ângulo de classe C^r. Mais precisamente, dado $\theta_0 \in \mathbb{R}$ tal que $z(a) = (\cos\theta_0, \operatorname{sen}\theta_0)$, z admite uma única função-ângulo $\theta\colon [a, b] \to \mathbb{R}$, tal que $\theta(a) = \theta_0$.*

Demonstração: Primeiro a unicidade (que demonstraremos mais geralmente, para funções-ângulo contínuas). Sejam $\theta, \varphi \colon [a,b] \to \mathbb{R}$ duas funções contínuas tais que $\xi \circ \theta = \xi \circ \varphi = z$. Para cada $t \in [a,b]$, $\theta(t) - \varphi(t)$ é um múltiplo inteiro de 2π. O inteiro $[\theta(t) - \varphi(t)]/2\pi$, dependendo continuamente de $t \in [a,b]$, é constante. Logo $\theta(a) = \varphi(a)$ implica $\theta(t) = \varphi(t)$ para todo $t \in [a,b]$.

Agora a existência de θ. Temos $z(t) = (x(t), y(t))$, onde $x, y \colon [a,b] \to \mathbb{R}$ são funções de classe C^r, com $x(a) = \cos\theta_0$, $y(a) = \operatorname{sen}\theta_0$. Como $|z(t)| = 1$ para todo t, o vetor $z' = (x', y')$ é, para todo valor de t, perpendicular a z, logo é um múltiplo escalar do vetor $w = (-y, x)$. Assim, para cada $t \in [a,b]$, existe um número $\lambda = \lambda(t)$ tal que $x' = -\lambda y$ e $y' = \lambda x$. Note que a função $\lambda \colon [a,b] \to \mathbb{R}$, assim definida, é de classe C^{r-1} pois $\lambda = \langle w, z' \rangle$. A função $\theta \colon [a,b] \to \mathbb{R}$ que procuramos é obtida pondo-se $\theta(t) = \theta_0 + \int_a^t \lambda(s)\,ds$. Vemos que θ é de classe C^r, com $\theta(a) = \theta_0$ e $\theta' = \lambda$. Devemos, em seguida, provar que $x(t) = \cos\theta(t)$ e $y(t) = \operatorname{sen}\theta(t)$ para todo $t \in [a,b]$. Ora, as relações $\theta' = \lambda$, $x' = -\lambda y$, $y' = \lambda x$ mostram que as funções $x\cos\theta + y\operatorname{sen}\theta$ e $y\cos\theta - x\operatorname{sen}\theta$ têm derivada nula, logo são constantes. Para $t = a$, a primeira é igual a 1 e a segunda é zero. Resolvendo o sistema de equações lineares

$$\begin{cases} x\cos\theta + y\operatorname{sen}\theta = 1 \\ -x\operatorname{sen}\theta + y\cos\theta = 0 \end{cases}$$

em relação às incógnitas x, y, obtemos $x = \cos\theta$ e $y = \operatorname{sen}\theta$.

Corolário 1. *Seja $f \colon [a,b] \to \mathbb{R}^2 - \{0\}$ um caminho de classe C^r $(r \geq 1)$. Dado $\theta_0 \in \mathbb{R}$ tal que $f(a) = |f(a)|(\cos\theta_0, \operatorname{sen}\theta_0)$, existe uma única função de classe C^r, $\theta \colon [a,b] \to \mathbb{R}$, tal que $\theta(a) = \theta_0$ e $f(t) = |f(t)|(\cos\theta(t), \operatorname{sen}\theta(t))$ para todo $t \in [a,b]$.*

Basta tomar $\theta =$ função-ângulo do caminho $z(t) = f(t)/|f(t)|$ com $\theta(a) = \theta_0$.

Corolário 2. *Seja $f \colon [a,b] \to \mathbb{R}^2 - \{0\}$ um caminho de classe C^r por partes $(r \geq 1)$. Dado $\theta_0 \in \mathbb{R}$ tal que $f(a) = |f(a)|(\cos\theta_0, \operatorname{sen}\theta_0)$, existe uma única função de classe C^r por partes, $\theta \colon [a,b] \to \mathbb{R}$, tal que $\theta(a) = \theta_0$ e $f(t) = |f(t)| \cdot (\cos\theta(t), \operatorname{sen}\theta(t))$ para todo $t \in [a,b]$.*

Basta mostrar que o caminho $z(t) = f(t)/|f(t)|$, de classe C^r por partes em $[a,b]$, admite uma função-ângulo θ, de classe C^r por partes, com $\theta(a) = \theta_0$. Se $f|[a,t_1]$ é de classe C^r, o Teorema 11 permite definir

θ no intervalo $[a, t_1]$. Se f é de classe C^r em $[t_1, t_2]$, novamente o Teorema 11 fornece uma função-ângulo, C^r por partes, para o caminho z, no intervalo $[a, t_2]$. Prosseguindo analogamente, definimos $\theta\colon [a, b] \to \mathbb{R}$, função de classe C^r por partes, tal que $z(t) = (\cos\theta(t), \operatorname{sen}\theta(t))$ para todo $t \in [a, b]$ e $\theta(a) = \theta_0$. Isto prova o corolário, pois a unicidade é evidente.

Observações: 1) Seja $f\colon [a, b] \to \mathbb{R}^2 - \{0\}$ um caminho de classe C^r $(r \geq 1)$. Se uma função contínua $\theta\colon [a, b] \to \mathbb{R}$ é tal que $f(t) = |f(t)|(\cos\theta(t), \operatorname{sen}\theta(t))$ para todo $t \in [a, b]$, então θ é uma função de classe C^r. Com efeito, seja φ uma função-ângulo de classe C^r para o caminho $z(t) = f(t)/|f(t)|$, tal que $\varphi(a) = \theta(a)$. Como a unicidade no Teorema 11 foi provada para funções contínuas, temos $\varphi(t) = \theta(t)$ para todo $t \in [a, b]$, logo θ é de classe C^r. O Exercício 7.1, adiante, pede para mostrar que todo caminho contínuo $z\colon [a, b] \to S^1$ admite uma função-ângulo contínua.

2) Na demonstração do Teorema 11, temos $w = iz$, onde $i = \sqrt{-1}$. Logo $z' = \lambda \cdot iz$. Se $\theta\colon [a, b] \to \mathbb{R}$ é tal que $\theta(a) = \theta_0$ e $\theta' = \lambda$, então a função $f = e^{i\theta}$ cumpre $f(a) = z(a)$ e $f' = \lambda \cdot if$. Isto justifica a definição de θ e nos dá esperança de que seja $f(t) = z(t)$ para todo $t \in [a, b]$. Para provar a igualdade $f = z$, mostramos que o quociente $z/f = z \cdot e^{-i\theta}$, que é igual a 1 no ponto $t = a$, é constante porque tem derivada nula. Ora, $z \cdot e^{-i\theta} = (x + iy)(\cos\theta - i\operatorname{sen}\theta) = [x\cos\theta + y\operatorname{sen}\theta] + i[y\cos\theta - x\operatorname{sen}\theta]$. Isto justifica o argumento usado na demonstração do teorema.

Exercícios

§1.

1.1. Seja $f\colon I \to \mathbb{R}^n$ um caminho diferenciável. Se $a \in I$ é ponto de acumulação do conjunto $f^{-1}(v)$, para algum $v \in \mathbb{R}^n$, então $f'(a) = 0$.

1.2. Seja $f\colon I \to \mathbb{R}^n$ um caminho diferenciável, com $f'(a) \neq 0$ para algum $a \in I$. Se existem uma reta $L \subset \mathbb{R}^n$ e uma seqüência de números distintos $t_k \to a$ tais que $f(t_k) \in L$ então L é a reta tangente a f no ponto $f(a)$.

1.3. Seja $f\colon I \to \mathbb{R}^n$ um caminho diferenciável. Dados $a \in \mathbb{R}^n$ e $r > 0$, a fim de que $f(t)$ pertença, para todo $t \in I$, à esfera de centro a e raio r, é necessário e suficiente que isto ocorra para um valor $t_0 \in I$ e que o vetor velocidade $f'(t)$ seja perpendicular a $f(t) - a$, para todo $t \in I$.

1.4. Seja $\lambda\colon [a, b] \to \mathbb{R}^n$ um caminho fechado diferenciável. Mostre que existe algum $t \in (a, b)$ tal que $\langle \lambda(t), \lambda'(t) \rangle = 0$.

1.5. Sejam $f_1, \ldots, f_p : I \to \mathbb{R}^m$ caminhos diferenciáveis e $\varphi : \mathbb{R}^m \times \cdots \times \mathbb{R}^m \to \mathbb{R}^n$ uma aplicação p-linear. Mostre que o caminho $g : I \to \mathbb{R}^n$, dado por $g(t) = \varphi(f_1(t), \ldots, f_p(t))$ é diferenciável e $g'(t) = \sum_{i=1}^{p} \varphi(f_1(t), \ldots, f_i'(t), \ldots, f_p(t))$. Conclua que se $f : (-\varepsilon, \varepsilon) \to \mathbb{R}^{n^2}$ é um caminho diferenciável, com $f(0) = I$ e $g : (-\varepsilon, \varepsilon) \to \mathbb{R}$ é definida por $g(t) = \det . f(t)$, então $g'(0) = \operatorname{tr}.A$ (traço de A), onde $A = f'(0)$.

1.6. Seja $X : I \to \mathbb{R}^{n^2}$ um caminho diferenciável cujos valores são matrizes $n \times n$. Fixado $k \in \mathbb{N}$, seja $f : I \to \mathbb{R}^{n^2}$ o caminho dado pela regra $f(t) = X(t)^k$. Mostre que f é diferenciável e determine $f'(t)$, para todo $t \in I$.

1.7. Se uma seqüência de caminhos $f_k : [a, b) \to \mathbb{R}^n$, deriváveis à direita, converge num ponto $x_0 \in [a, b)$ e suas derivadas à direita $f'_{k+} : [a, b) \to \mathbb{R}^n$ convergem uniformemente para um caminho g, então (f_k) converge uniformemente para um caminho f, derivável à direita, e vale $f'_+ = g$.

1.8. Para toda matriz $X \in \mathbb{R}^{n^2}$, seja $e^X = I + X + \dfrac{X^2}{2} + \cdots + \dfrac{X^n}{n!} + \ldots$. Defina um caminho $f : \mathbb{R} \to \mathbb{R}^{n^2}$ pondo $f(t) = e^{tA}$, onde $A \in \mathbb{R}^{n^2}$ é fixada. Mostre que $f'(0) = A$.

1.9. Estenda o Teorema de Ascoli-Arzelá (vol. 1, pág. 412) para caminhos contínuos $f : [a, b] \to \mathbb{R}^n$.

1.10. Seja $f : I \to \mathbb{R}^n$ um caminho diferenciável, com $f'(a) \neq 0$ para um certo $a \in I$. Uma reta $L \subset R^n$, contendo o ponto $f(a)$, é a reta tangente a f nesse ponto se, e somente se,

$$\lim_{t \to a} \frac{d(f(t), L)}{|f(t) - f(a)|} = 0.$$

1.11. Sejam $f : [a, b) \to \mathbb{R}^2$ um caminho (admite-se $b = +\infty$) tal que $\lim_{t \to b} |f(t)| = \infty$ e $L = \{(x, y) \in \mathbb{R}^2 ; \alpha x + \beta y = c\}$ uma reta. Ponhamos $u = (\alpha, \beta)$. Podemos supor $|u|^2 = \alpha^2 + \beta^2 = 1$. As seguintes afirmações são equivalentes:

(i) $\lim_{t \to b} d(f(t), L) = 0$;

(ii) $\lim_{t \to b} \langle f(t), u \rangle = c$ e $\lim_{t \to b} \left\langle \dfrac{f(t)}{|f(t)|}, u \right\rangle = 0$.

Quando isto ocorre, diz-se que a reta L é *assíntota* do caminho f quando $x \to b$.

1.12. Se $b < +\infty$ e o caminho $f : [a, b) \to \mathbb{R}^2$ é da forma $f(t) = (t, \varphi(t))$, com $\lim_{t \to b} \varphi(t) = +\infty$, a reta vertical $x = b$ é assíntota do caminho f quando $t \to b$.

1.13. Mostre que uma parábola $f : \mathbb{R} \to \mathbb{R}^2$, $f(t) = (t, at^2 + bt + c)$, não possui assíntota. Por outro lado, determine a assíntota da hipérbole $f : [a, +\infty) \to \mathbb{R}^2$, $f(t) = (t, \sqrt{t^2 - a})$.

1.14. Seja $f : (a, b) \to \mathbb{R}^n$ um caminho contínuo que possui em cada ponto de (a, b) um vetor velocidade à direita. Se $f'_+ : (a, b) \to \mathbb{R}^n$ é contínuo, prove que f é de classe C^1.

1.15. Seja $G \subset \mathbb{R}^2$ o gráfico da função $y = |x|$. Mostre que existe um caminho $f\colon \mathbb{R} \to \mathbb{R}^2$, de classe C^∞, tal que $f(\mathbb{R}) = G$ e $f(0) = (0,0)$. Prove que todo caminho diferenciável f nessas condições cumpre $f'(0) = 0$.

1.16. Seja $f\colon (a,b) \to \mathbb{R}^n$ um caminho diferenciável. Defina $\varphi\colon (a,b) \to \mathbb{R}$ pondo $\varphi(t) = |f(t) - p|$, onde p é um ponto de $\mathbb{R}^n$ que não pertence à imagem de f. (Norma euclidiana.) Prove que $|\varphi'(t)|$ é o comprimento da projeção de $f'(t)$ sobre o eixo $\overline{pf(t)}$.

1.17. Seja $f\colon (a,b) \to \mathbb{R}^n$ um caminho diferenciável, com $f(t) \neq 0$ para todo t. Calcule o vetor velocidade do caminho $\varphi\colon (a,b) \to \mathbb{R}^n$, definido por $\varphi(t) = \dfrac{f(t)}{|f(t)|}$ (norma euclidiana). No caso particular de $f\colon \mathbb{R} \to \mathbb{R}^n$, dado por $f(t) = u + th$, com $u \in S^{n-1}$ e $h \in \mathbb{R}^n$, calcule $\varphi'(0)$ e conclua que h é perpendicular ao vetor unitário u se, e somente se, é o vetor velocidade $h = \varphi'(0)$ de um caminho $\varphi\colon \mathbb{R} \to S^{n-1}$ com $\varphi(0) = u$.

1.18. Seja $f\colon I \to \mathbb{R}^n$ um caminho diferenciável tal que $f(t) \neq 0$ e $f'(t) = \lambda(t) \cdot f(t)$ para todo $t \in I$. Prove que a imagem $f(I)$ está contida numa reta que passa pela origem. $\left[\text{Derive } \varphi(t) = \dfrac{f(t)}{|f(t)|} \cdot \right]$

1.19. Use o exercício anterior para mostrar que se f é diferenciável, com $f(t) = \lambda(t) \cdot f'(t)$ e $f'(t) \neq 0$ para todo t então $f(I)$ está contido numa reta que passa pela origem.

1.20. Seja (x_k) uma seqüência de pontos em $\mathbb{R}^m - \{0\}$ tal que $|x_k| > |x_{k+1}|$, $\lim x_k = 0$ e $\lim \dfrac{x_k}{|x_k|} = u \in S^{m-1}$. Mostre que existe um caminho $f\colon \mathbb{R} \to \mathbb{R}^m$, diferenciável no ponto $t = 0$, com $f'(0) = u$, e uma seqüência de números $t_k > 0$ tais que $\lim t_k = 0$ e $f(t_k) = x_k$.

1.21. Seja $f\colon (-\varepsilon, \varepsilon) \to \mathbb{R}^{n^2}$ diferenciável no ponto $t = 0$, com $f(0) = I =$ matriz identidade $n \times n$ e $f'(0) = A$. Se $f(t)$ é, para todo t, uma matriz ortogonal, então A é anti-simétrica. Se $\det .f(t) = 1$ para todo t então o traço de A é zero.

1.22. Sejam $A, B \in \mathbb{R}^{n^2}$, A anti-simétrica e B com traço nulo. Mostre que, para todo $t \in \mathbb{R}$, e^{tA} é ortogonal e e^{tB} tem determinante 1. Conclua que toda matriz anti-simétrica é vetor velocidade de um caminho de matrizes ortogonais e toda matriz de traço nulo é vetor velocidade de um caminho de matrizes unimodulares ($\det = 1$).

§2.

2.1. Se $f, g\colon [a,b] \to \mathbb{R}^n$ são de classe C^1 então

$$\int_a^b \langle f(t), g'(t)\rangle\, dt = \langle f, g\rangle\Big|_a^b - \int_a^b \langle f'(t), g(t)\rangle\, dt.$$

2.2. Se uma seqüência de caminhos integráveis $f_k\colon [a,b] \to \mathbb{R}^n$ converge uniformemente para um caminho $f\colon [a,b] \to R^n$ então f é integrável e

$$\lim_{l\to\infty} \int_a^b f_k(t)\, dt = \int_a^b f(t)\, dt.$$

2.3. Seja $A \subset \mathbb{R}^m$ um conjunto convexo. Dado um caminho integrável $f\colon [0,1] \to \mathbb{R}^m$ tal que $f(t) \in A$ para todo t, prove que $\int_0^1 f(t)\,dt \in \overline{A}$.

2.4. Prove, diretamente a partir da definição dada no texto, que se os caminhos $f, g\colon [a,b] \to \mathbb{R}^n$ são integráveis e $\alpha, \beta \in \mathbb{R}$, então o caminho $\alpha f + \beta g$ é integrável e $\int_a^b (\alpha f + \beta g)\,dt = \alpha \int_a^b f\,dt + \beta \int_a^b g\,dt$.

2.5. Seja $\times$ o produto vetorial em $\mathbb{R}^3$. Dados um vetor $v \in \mathbb{R}^3$ e um caminho integrável $f\colon [a,b] \to \mathbb{R}^3$, mostre que $\int_a^b (v \times f(t))\,dt = v \times \int_a^b f(t)\,dt$.

§3.

3.1. Sejam $x\colon \mathbb{R} \to \mathbb{R}^n$ um caminho diferenciável e $A\colon \mathbb{R}^n \to \mathbb{R}^n$ uma transformação linear. Se $x'(t) = A \cdot x(t)$ para todo $t \in \mathbb{R}$, prove que $x \in C^\infty$. Mais ainda: o caminho x é analítico e, para todo $a \in \mathbb{R}$, tem-se $x(t) = e^{(t-a)A} \cdot x(a)$. Em particular, se $x(a) = 0$ para algum $a \in \mathbb{R}$ então $x(t) = 0$ para todo $t \in \mathbb{R}$. Ou melhor: se $x, y\colon \mathbb{R} \to \mathbb{R}^n$ são diferenciáveis com $x' = A \cdot x$, $y' = A \cdot y$ e $x(a) = y(a)$ para algum $a \in \mathbb{R}$, então $x = y$. [Isto é bem mais fácil do que parece. Basta derivar sucessivamente a igualdade $x' = A \cdot x$ para obter a série de Taylor de $x(t)$.]

3.2. Seja $A\colon \mathbb{R}^2 \to \mathbb{R}^2$ dada por $A \cdot (x,y) = (-cy, cx)$. Um caminho duas vezes diferenciável $X\colon \mathbb{R} \to \mathbb{R}^2$ cumpre $X' = A \cdot X$ e $X(0) = (a,b)$ se, e somente se, $X(t) = (x(t), y(t))$, com $x' = -cy$, $x'' + c^2 x = 0$, $x(0) = a$ e $y(0) = b$. Compute e^{tA} e conclua que a única função duas vezes diferenciável $x\colon \mathbb{R} \to \mathbb{R}$ tal que $x'' + c^2 x = 0$, $x(0) = a$ e $x'(0) = b$ é a função $x(t) = a \cos ct + b \operatorname{sen} ct$. Conclua ainda que se $F\colon \mathbb{R} \to \mathbb{R}^n$ é um caminho duas vezes diferenciável tal que $F'' + c^2 F = 0$ então existem vetores $A, B \in \mathbb{R}^n$ tais que $F(t) = \cos ct \cdot A + \operatorname{sen} ct \cdot B$ para todo $t \in R$. Em particular, F é um caminho plano.

3.3. A partir de definição de vetor velocidade, usando o argumento de bisseções repetidas (como no Lema 2 do §3) prove que se $f\colon [a,b] \to \mathbb{R}^n$ é um caminho diferenciável com $f'(t) = 0$ para todo t então f é constante.

3.4. Sejam $f\colon [a,b] \to \mathbb{R}^n$, $g\colon [a,b] \to \mathbb{R}$ contínuas, deriváveis à direita em (a,b). Se $|f'_+(x)| \le g'_+(x)$ para todo $x \in (a,b)$ então $|f(b) - f(a)| \le g(b) - g(a)$. Conclua que toda função real contínua num intervalo fechado que possui derivada à direita ≥ 0 em todos os pontos interiores ao intervalo é não-decrescente. Conclua também que se $|f'_+(x)| \le c$ par todo $x \in (a,b)$ então o caminho contínuo $f\colon [a,b] \to \mathbb{R}^n$ cumpre $|f(x) - f(y)| \le c|x - y|$ para quaisquer $x, y \in [a,b]$.

3.5. Seja $f\colon (a,b) \to \mathbb{R}^n$ tal que, para um certo $c \in (a,b)$, as restrições $f|(a,c]$ e $f|[c,b)$ são de classe C^k $(1 \le k \le \infty)$, podendo ser $f'(c^-) \ne f'(c^+)$. Obtenha um homeomorfismo $\varphi\colon (\alpha,\beta) \to (a,b)$, de classe C^∞, com $\varphi'(t) > 0$ exceto num ponto γ, de modo que $f \circ \varphi\colon (\alpha,\beta) \to \mathbb{R}^n$ seja de classe C^k.
[Sugestão: todas as derivadas de φ devem ser zero no ponto γ.] Conclua que dois pontos quaisquer de um aberto conexo em $\mathbb{R}^n$ podem ser ligados por um caminho C^∞.

§4.

4.1. Sejam $f\colon [0,2\pi] \to \mathbb{R}$ e $g\colon [0,2\pi] \to \mathbb{R}^2$ definidos por $f(t) = \operatorname{sen} t$ e $g(t) = (t, \operatorname{sen} t)$. Determine $\ell(f)$ e $\ell(g)$.

4.2. Qual o comprimento da ciclóide $f\colon [0,2\pi] \to \mathbb{R}^2$, dada por $f(t) = (t - \text{sen } t, 1 - \cos t)$?

4.3. Seja $U \subset \mathbb{R}^m$ aberto e conexo. Dados $x, y \in U$, defina a *distância geodésica* $d_U(x,y)$ como o ínfimo dos comprimentos dos caminhos contínuos retificáveis contidos em U, que ligam x a y.

(a) $d_U(x,y)$ é o ínfimo dos comprimentos dos caminhos poligonais contidos em U que ligam x a y.

(b) Tem-se $d_U(x,z) \le d_U(x,y) + d_U(y,z)$ e $d_U(x,y) = 0 \Rightarrow x = y$.

(c) Dada uma seqüência de pontos $x_k \in U$ vale $\lim\limits_{k\to\infty} |x_k - a| = 0 \quad (a \in U)$ se, e somente se, $\lim\limits_{k\to\infty} d_U(x_k, a) = 0$.

(d) Dê exemplo em que se tenha $\lim\limits_{k\to\infty} |x_k - y_k| = 0$, com $y_k, x_k \in U$ mas $\lim\limits_{k\to\infty} d_U(x_k, y_k) \neq 0$.

(e) Se $V \subset \mathbb{R}^m$ é convexo e $U = V - F$, onde F é finito, então $d_U(x,y) = |x - y|$ para qualquer $x, y \in U$.

(f) Suponha que dois pontos quaisquer de $U \subset \mathbb{R}^m$ podem ser ligados por um caminho contínuo retificável contido em U, mas não exija que U seja aberto. Então $d_U(x,y)$ pode ainda ser definida e vale (b) mas as propriedades (a) e (c) podem falhar. Se U é a esfera unitária S^{m-1}, mostre que $d_U(x,y)$ satisfaz (b) e (c).

4.4. Dada uma seqüência de caminhos retificáveis $f_k\colon [a,b] \to \mathbb{R}^n$, suponha que exista $c > 0$ tal que $\ell(f_k) \le c$ para todo k. Se (f_k) converge simplesmente para o caminho $f\colon [a,b] \to \mathbb{R}^n$, prove que f é retificável e $\ell(f) \le c$.

4.5. Nas condições do exercício anterior, mostre que $\lim\limits_{k\to\infty} \ell(f_k)$ pode não existir e, mesmo que exista $L = \lim \ell(f_k)$, tem-se apenas $\ell(f) \le L$, sem que valha necessariamente a igualdade.

4.6. Considere a seqüência de caminhos $f_n\colon [0,2\pi] \to \mathbb{R}^2$, onde $f_n(t) = (t, \text{sen}(nt)/n)$. Mostre que $f_n \to f$ uniformemente, onde $f(t) = (t, 0)$, mas $\lim\limits_{n\to\infty} \ell(f_n) > \ell(f)$.

4.7. Sejam $a, b \in K$, onde $K \subset \mathbb{R}^n$ é compacto. Seja C o conjunto dos caminhos contínuos $f\colon [0,1] \to K$, retificáveis, com $f(0) = a$ e $f(1) = b$. Se $C \neq \varnothing$ então existe $f_0 \in C$ tal que $\ell(f_0) \le \ell(f)$ para todo $f \in C$. [Sugestão: O conjunto X dos caminhos $f \in C$ com $\ell(f) \le k$ e $f = g \circ \varphi$, onde $g\colon [0,L] \to K$ é cadenciado e $\varphi\colon [0,1] \to [0,L]$ é da forma $\varphi(t) = L \cdot t$, é equicontínuo e equilimitado, pois $f \in X \Rightarrow |f(s) - f(t)| \le k|s - t|$. Use Ascoli-Arzelá.]

4.8. Sejam $f, g\colon [a,b] \to \mathbb{R}^n$ caminhos tais que f é retificável e o conjunto $F = \{t \in [a,b]; f(t) \neq g(t)\}$ é finito. Mostre que g é retificável. Mais geralmente, se $F = \{t_1, \ldots, t_i, \ldots\}$ é enumerável e, além disso, $\sum\limits_{i=1}^{\infty} |f(t_i) - g(t_i)| < +\infty$ então g é retificável.

4.9. Todo caminho retificável $f\colon [a,b] \to \mathbb{R}^n$ é integrável.

4.10. Sejam $c_1, \ldots, c_i, \ldots$ os pontos de descontinuidade de um caminho retificável $f\colon [a,b] \to \mathbb{R}^n$. Ponha $u_i = \lim\limits_{t\to c_i^-} f(t)$, $v_i = \lim\limits_{t\to c_i^+} f(t)$ e $s_i = |f(c_i) - u_i| + |f(c_i) + v_i|$. Prove que a série $\Sigma\, s_i$ é convergente.

4.11. Seja $f\colon [0,1] \to \mathbb{R}$ definida por $f(t) = 0$ se t é irracional e $f(p/q) = 1/q$ se $t = p/q$ é uma fração irredutível. Prove que f não é retificável (de variação limitada), embora possua limites laterais em todos os pontos.

4.12. Seja $f\colon [a,b] \to \mathbb{R}^n$ um caminho diferenciável. Se a função $|f'|\colon [a,b] \to \mathbb{R}$ for integrável (em particular, se $f'\colon [a,b] \to \mathbb{R}^n$ for integrável) então f é retificável e vale $\ell(f) = \int_a^b |f'(t)|\,dt$. A retificabilidade de f resulta da condição de Lipschitz, que é satisfeita porque f' é limitada. A fórmula para o comprimento de arco decorre do seguinte "*Lema*: Para todo $n \in \mathbb{N}$, existe uma partição pontilhada $P_n^* = (P_n, \xi)$ de $[a,b]$, tal que $|P_n| < 1/n$ e $|f(t_i) - f(t_{i-1}) - f'(\xi_i)(t_i - t_{i-1})| < (t_i - t_{i-1})/n$, seja qual for o intervalo $[(t_{i-1}, t_i]$ da partição P_n." A demonstração do lema se faz sem dificuldade, usando o método das bisseções. Dele resulta que $|\ell(f; P_n) - \Sigma(|f'|; P_n^*)| < (b-a)/n$. Fazendo $n \to \infty$ obtém-se a fórmula procurada.

§5.

5.1. Dado o caminho retificável $f\colon [a,b] \to \mathbb{R}^n$, defina $\sigma\colon [a,b] \to \mathbb{R}$ pondo $\sigma(t) = \ell(f|[a,t])$. Se σ é uma função contínua, então f também o é.

5.2. Seja $f\colon I \to \mathbb{R}^n$ um caminho diferenciável. Dada uma função contínua positiva $\varphi\colon X \to \mathbb{R}$, cujo domínio $X \subset \mathbb{R}^n$ contenha todos os pontos $f(t)$, $t \in I$, podemos reparametrizar f por meio de um difeomorfismo C^1, de modo que o vetor velocidade em cada ponto $p = f(t)$ seja agora $\varphi(p) \cdot f'(t)$, em vez de $f'(t)$.

5.3. Defina $f\colon [0,1] \to \mathbb{R}$ pondo $f(0) = 0$ e $f(t) = t^2 \operatorname{sen}(1/t^2)$ quando $t \neq 0$. Prove que f é um caminho diferenciável porém não retificável.

§6.

6.1. Seja $f\colon \mathbb{R} \to \mathbb{R}^3$ a hélice, dada por $f(t) = (a\cos bt, a \operatorname{sen} bt, ct)$. Determine a relação entre a, b e c para que o caminho f seja cadenciado. Supondo isto satisfeito, mostre que a curvatura e a torção de f são constantes.

6.2. Seja $f\colon I \to \mathbb{R}^3$ um caminho cadenciado de classe C^2. Se, para todo $x \in I$, a reta normal $L = \{f(s) + t \cdot f''(s); t \in \mathbb{R}\}$ contém a origem então imagem $f(I)$ está contida num plano.

6.3. Seja $f\colon I \to \mathbb{R}$ um caminho regular de classe C^2, não necessariamente cadenciado. A curvatura $k(t)$ de f num ponto $p = f(t)$ é, por definição, a curvatura do caminho cadenciado $g = f \circ \varphi$, obtido por reparametrização de f, no mesmo ponto $p = g(s)$. Mostre que

$$k(t) = \frac{|f'(t) \times f''(t)|}{|f'(t)|^3}\,.$$

Se $f\colon I \to \mathbb{R}^2$ é um caminho plano, mostre que se tem

$$k(t) = \frac{\det[f'(t), f''(t)]}{|f'(t)|^3}\,.$$

Considere o caso particular em que $f(t) = (t, \varphi(t))$.

6.4. Seja $f\colon I \to \mathbb{R}^3$ um caminho regular de classe C^3. Definindo a torção $w(t)$ de modo análogo ao do exercício anterior, mostre que

$$w(t) = \frac{\det[f'(t), f''(t), f'''(t)]}{|f'(t) \times f''(t)|^2}\,.$$

6.5. Seja $f\colon I \to \mathbb{R}^3$ um caminho cadenciado de classe C^3. Se todos os planos osculadores de f são paralelos, (ou contêm um ponto dado, digamos, a origem) então a imagem $f(I)$ está contida num plano.

6.6. Seja $f\colon I \to \mathbb{R}^n$ um caminho cadenciado de classe C^n. Para cada $s \in I$ defina o "n-edro de Frenet"$\{e_1, \dots, e_n\}$, $e_i = e_i(s)$ e as curvaturas $k_1, \dots, k_n$, $k_i = k_i(s)$, pondo $e_i = f'$ e, supondo $e_1' \neq 0$, escrevendo $e_1' = k_1 \cdot e_2$ com $|e_2| = 1$ e $e_2 \perp e_1$. Em seguida, notando que $e_2' \perp e_2$ e $\langle e_2', e_1\rangle = -\langle e_2, e_1'\rangle = -k_1$, escreva $k_2 = |k_1 e_1 + e_2'|$. Se $k_2 > 0$, ponha $e_2' = -k_1 e_1 + k_2 e_3$, o que define univocamente o triedro ortonormal $\{e_1, e_2, e_3\}$ e a segunda curvatura k_2. Note que $\langle e_3', e_1\rangle = -\langle e_3, e_1'\rangle = 0$, logo $e_3' \perp e_3$, $e_3' \perp e_1$ e $\langle e_3', e_2\rangle = -\langle e_3, e_2'\rangle = -k_2$. Ponha $k_3 = |e_3' + k_2 e_3|$. Se $k_3 > 0$, o vetor unitário e_4 fica bem definido pela igualdade $e_3' = -k_2 e_2 + k_3 e_4$ de modo que $\{e_1, e_2, e_3, e_4\}$ é um conjunto ortonormal. Prossiga.

6.7. Mostre que dois caminhos planos cadenciados de classe C^2, são congruentes se, e somente se, têm o mesmo comprimento e a mesma função curvatura. Mais precisamente, sejam $f, g\colon [0, L] \to \mathbb{R}^2$ de classe C^2 tais que $|f'(s)| = |g'(s)| = 1$ e $k_f(s) = k_g(s)$ para todo $s \in [0, L]$. Prove que existe uma isometria $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ tal que $g(s) = T[f(s)]$ para todo $s \in [0, L]$.

6.8. Seja $f\colon I \to \mathbb{R}^n$ um caminho regular de classe C^2. Mostre que, em cada ponto $f(t)$, a aceleração $f''(t)$ se decompõe na soma $f'' = v' \cdot T + kv^2 \cdot N$, onde $T = f'/|f|$, $N = T'/|T'|$, $v = |f'|$ é a velocidade escalar e k é a curvatura. [Que tem isto a ver com a afirmação de que é vantajoso acelerar um carro no meio de uma curva?]

§7.

7.1 Sejam $f\colon [a, b] \to \mathbb{R}^2 - \{0\}$, $f \in C^0$, $\theta_0 \in \mathbb{R}$ com $f(a) = |f(a)|(\cos\theta_0, \operatorname{sen}\theta_0)$, existe uma (única) função contínua $\theta\colon [a, b] \to \mathbb{R}$ tal que $\theta(a) = \theta_0$ e, para todo $t \in [a, b]$, $f(t) = |f(t)|(\cos\theta(t), \operatorname{sen}\theta(t))$. [Sugestão: Basta considerar o caso em que $f([a, b]) \subset S^1$, ou seja, $|f(t)| = 1$ para todo t. Suponha inicialmente que exista $c \in S^1 - f([a, b])$, considere um homeomorfismo $h\colon (\alpha, \alpha + 2\pi) \to S^1 - \{c\}$ e defina $\theta = h^{-1} \circ f$. Reduza o caso geral a este subdividindo o intervalo $[a, b]$ em um número finito de subintervalos tais que em cada um deles f não seja sobrejetiva.]

Capítulo III

Funções Reais de n Variáveis

1 Derivadas parciais

Quando se estudam funções reais de n variáveis, isto é, definidas em subconjuntos do espaço $\mathbb{R}^n$, e se busca para essas funções uma noção de derivada que tenha propriedades análogas às da derivada de uma função definida num intervalo, a idéia que se apresenta mais naturalmente é a de "derivada parcial", que exporemos agora.

Para efeito de derivação, onde se compara o acréscimo $f(a+h) - f(a)$ da função f com o acréscimo h dado ao ponto a, o domínio mais adequado para uma função é um subconjunto aberto $U \subset \mathbb{R}^n$ pois, neste caso, dado $a \in U$, tem-se ainda $a + h \in U$ para todo acréscimo suficientemente pequeno h. Seja, pois, $f\colon U \to \mathbb{R}$ uma função real, definida num subconjunto aberto $U \subset \mathbb{R}^n$. Dado o ponto $a \in U$, a *i-ésima derivada parcial* de f no ponto a (onde $1 \leq i \leq n$) é o limite

$$\frac{\partial f}{\partial x_i}(a) = \lim_{t \to 0} \frac{f(a + te_i) - f(a)}{t}\,,$$

quando tal limite existe. Às vezes, usaremos também a notação $\partial_i f(a)$.

Observação: O símbolo $\dfrac{\partial f}{\partial x_i}$ terá para nós o mesmo significado que $\dfrac{\partial f}{\partial y_i}$, $\dfrac{\partial f}{\partial z_i}$, etc. O que importa num símbolo destes não é o "nome" da variável,

que tanto pode ser x, como y ou z, etc. O importante é o índice i: trata-se da derivada de f em relação à sua i-ésima variável, seja qual for o sinal usado pra indicá-la. Estritamente falando, a melhor notação para a i-ésima derivada parcial é $\partial_i f$, mas continuaremos escrevendo $\dfrac{\partial f}{\partial x_i}$ por respeito à tradição, pelo apelo estético e, principalmente, porque isto torna mais naturais certas fórmulas, como por exemplo a Regra da Cadeia.

Quando $U \subset \mathbb{R}^2$, uma função $f\colon U \to \mathbb{R}$ é o que se chama uma "função real de duas variáveis reais". Escreve-se $f(x, y)$ para indicar seu valor no ponto $z = (x, y)$. Desta forma, as derivadas parciais de f num ponto $c = (a, b) \in U$ podem também ser representadas por $\dfrac{\partial f}{\partial x}(c)$ e $\dfrac{\partial f}{\partial y}(c)$, em vez de $\dfrac{\partial f}{\partial x_1}(c)$, $\dfrac{\partial f}{\partial x_2}(c)$. Temos:

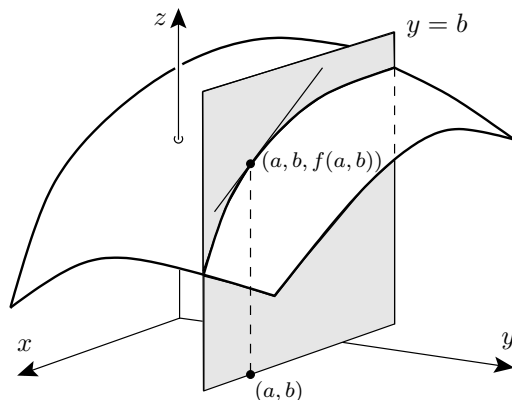
$$\frac{\partial f}{\partial x}(c) = \lim_{t\to 0} \frac{f(a+t, b) - f(a, b)}{t},$$
$$\frac{\partial f}{\partial y}(c) = \lim_{t\to 0} \frac{f(a, b+t) - f(a, b)}{t}.$$

Analogamente, se $U \subset \mathbb{R}^3$, uma função $f\colon U \to \mathbb{R}$ é uma "função real de três variáveis reais". Seu valor num ponto $p = (x, y, z)$ se escreve $f(x, y, z)$ e suas derivadas parciais no ponto $q = (a, b, c)$ podem ser escritas como $\dfrac{\partial f}{\partial x}(q)$, $\dfrac{\partial f}{\partial y}(q)$ e $\dfrac{\partial f}{\partial z}(q)$.

Voltando ao caso geral, seja $f\colon U \to \mathbb{R}$ definida no aberto $U \subset \mathbb{R}^n$. Dados o ponto $a \in U$ e o inteiro $i \in [1, n]$, a imagem do caminho $\lambda\colon \mathbb{R} \to \mathbb{R}^n$, $\lambda(t) = a + te_i$, é o que se chama "a reta que passa por a e é paralela ao i-ésimo eixo". (Note que $\lambda(0) = a$.) Como U é aberto, existe $\varepsilon > 0$ tal que $-\varepsilon < t < \varepsilon \Rightarrow \lambda(t) = a + te_i \in U$. A i-ésima derivada parcial de f no ponto a é a derivada, no ponto $t = 0$, da função $f \circ \lambda\colon (-\varepsilon, \varepsilon) \to \mathbb{R}$, ou seja, $\dfrac{\partial f}{\partial x_i}(a) = (f \circ \lambda)'(0)$. Podemos dizer que f, quando restrita ao segmento de reta aberto $J = (a - \varepsilon e_i, a + \varepsilon e_i)$, torna-se uma função real, $f(a + te_i)$, da variável real t e $\dfrac{\partial f}{\partial x_i}(a)$ é a derivada dessa função no ponto $t = 0$.

Quando $n = 2$, o gráfico de f é uma superfície em $\mathbb{R}^3$; a restrição de f ao segmento de reta que passa por $c = (a, b)$ e é paralelo ao eixo das abcissas tem como gráfico a curva plana obtida nessa superfície fazendo

y constante, igual a b. Logo $(\partial f/\partial x)(c)$ é a inclinação da reta tangente a esssa curva, no ponto $(a, b, f(a, b))$, relativamente ao plano horizontal.

O cálculo prático da i-ésima derivada parcial de uma função real $f(x_1, \ldots, x_n)$ se faz considerando todas as variáveis como se fossem constantes, exceto a i-ésima, e aplicando as regras usuais de derivação relativamente a essa variável.

O comportamento da i-ésima derivada parcial $\partial f/\partial x_i$ ao longo de um segmento de reta contido no domínio de f e paralelo ao i-ésimo eixo dá informações sobre o crescimento de f ao longo de tal segmento. Assim, por exemplo, se $f\colon U \to \mathbb{R}$ está definida em $U \subset \mathbb{R}^2$, se o segmento de reta $J = \{(a, t); 0 \leq t \leq 1\}$, paralelo ao eixo dos y, está contido em U e, além disso $(\partial f/\partial y)(z) > 0$ para todo $z \in J$, então f é crescente sobre J, isto é, $0 \leq s < t \leq 1 \Rightarrow f(a, s) < f(a, t)$.

Ainda nessa ordem de idéias, dizemos que uma função $f\colon U \to \mathbb{R}$ *não depende da i-ésima variável* quando, dados $a = (a_1, \ldots, a_{i-1}, x, a_{i+1}, \ldots, a_n)$ e $b = (a_1, \ldots, a_{i-1}, y, a_{i+1}, \ldots, a_n)$ em U, tem-se $f(a) = f(b)$. Noutros termos, se $a, b \in U$ com $b = a + te_i$ então $f(a) = f(b)$.

Um conjunto $U \subset \mathbb{R}^n$ diz-se *i-convexo* quando $a, b \in U$, $\quad b = a + te_i \Rightarrow [a, b] \subset U$.

A afirmação abaixo resulta das definições: *Sejam $U \subset \mathbb{R}^n$ um aberto i-convexo e $f\colon U \to \mathbb{R}^n$ uma função tal que $\dfrac{\partial f}{\partial x_i}(x) = 0$ para todo $x \in U$. Então f é independente da i-ésima variável.*

Com efeito, se $a, b \in U$, com $b = a + te_i$, então a função $\xi\colon [0, t] \to \mathbb{R}$, definida por $\xi(s) = f(a + se_i)$, possui derivada $\xi'(s) = \dfrac{\partial f}{\partial x_i}(a + se_i) = 0$

para todo $s \in [0,t]$, logo é constante, e assim $f(a) = \xi(0) = \xi(t) = f(b)$.

No plano, diz-se *horizontalmente*, ou *verticalmente, convexo*, em vez de 1-convexo e 2-convexo, respectivamente.

Exemplo 1. Seja $X = \{(x,0) \in \mathbb{R}^2; x \geq 0\}$ o semi-eixo positivo fechado das abscissas. O aberto $U = \mathbb{R}^2 - X$ é horizontalmente (mas não verticalmente) convexo. A função $f\colon U \to \mathbb{R}$, definida por $f(x,y) = x^2$ quando $x > 0$, $y > 0$ e $f(x,y) = 0$ quando $x \leq 0$ ou $y \leq 0$, possui derivada parcial $\frac{\partial f}{\partial y} = 0$ em todos os pontos de U, mas f não é independente da segunda variável, y, pois se tomarmos $x > 0$, $y > 0$, teremos $f(x,y) = x^2 > 0$ e $f(x,-y) = 0$.

Note-se que as derivadas parciais sozinhas não permitem conclusões sobre o comportamento "n-dimensional" da função. Por exemplo, a existência de todas as derivadas parciais num ponto não implica a continuidade da função nesse ponto, como veremos agora.

Exemplo 2. Seja $f\colon \mathbb{R}^2 \to \mathbb{R}$ definida por $f(x,y) = xy/(x^2+y^2)$ se $x^2+y^2 \neq 0$ e $f(0,0) = 0$. Se $z = (x,y)$ não é a origem, temos $\partial_1 f(z) = (y^3 - x^2y)/(x^2+y^2)^2$ e $\partial_2 f(z) = (x^3 - xy^2)/(x^2+y^2)^2$. Na origem, vale:

$$\frac{\partial f}{\partial x}(0,0) = \lim_{t\to 0} \frac{f(t,0) - f(0,0)}{t} = 0 \quad \text{e}$$
$$\frac{\partial f}{\partial y}(0,0) = \lim_{t\to 0} \frac{f(0,t) - f(0,0)}{t} = 0.$$

Assim, f possui derivadas parciais em todos os pontos do plano. Entretanto, como vimos no Capítulo 1, f é descontínua na origem. [Se $x^2+y^2 \neq 0$, então $f(x,) = (x/\sqrt{x^2+y^2}) \cdot (y/\sqrt{x^2+y^2}) = \cos\theta \cdot \operatorname{sen}\theta$, onde θ é o ângulo formado pelo semi-eixo positivo das abcissas e a semireta que passa na origem e contém o ponto (x,y). Ao longo de cada uma dessas semi-retas, $f(x,y)$ tem valor constante, dependendo da semi-reta, logo não existe $\lim f(x,y)$ na origem.]

2 Derivadas direcionais

Vendo que as derivadas parciais, desacompanhadas de hipóteses adicionais, apenas fornecem informações sobre a função ao longo de retas paralelas aos eixos, tentamos estender a noção de derivada a outras

direções além dessas. Isto nos leva ao importante conceito de derivada direcional.

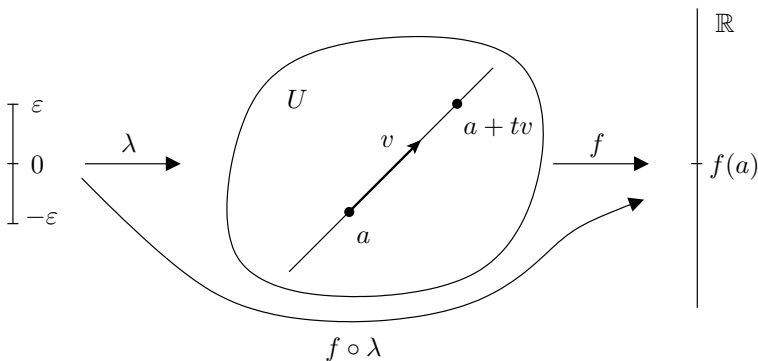
Sejam $f\colon U \to \mathbb{R}$ definida no aberto $U \subset \mathbb{R}^n$, $a \in U$ e $v \in \mathbb{R}^n$. A *derivada direcional* de f no ponto a, segundo o vetor v, é, por definição, o limite

$$\frac{\partial f}{\partial v}(a) = \lim_{t \to 0} \frac{f(a+tv) - f(a)}{t}$$

quando tal limite existe.

As derivadas parciais tornam-se casos particulares das derivadas direcionais:

$$\frac{\partial f}{\partial x_i}(a) = \frac{\partial f}{\partial e_i}(a) = \text{derivada direcional de } f \text{ no ponto } a,$$
$$\text{segundo o vetor } e_i\,.$$

A derivada direcional $(\partial f/\partial v)(a)$ é a derivada, no ponto $t = 0$, da função composta $f \circ \lambda\colon (-\varepsilon, \varepsilon) \to \mathbb{R}^n$, onde $\lambda\colon (-\varepsilon, \varepsilon) \to \mathbb{R}^n$ é o caminho retilíneo, $\lambda(t) = a + tv$, para o qual se tem $\lambda(0) = a$ e $\lambda'(t) = v$ para todo t. Aqui, $\varepsilon > 0$ é escolhido tão pequeno que a imagem de λ esteja contida em U.

A função f do Exemplo 2 possui as derivadas direcionais $\frac{\partial f}{\partial v}(0,0)$ para $v = (\alpha, 0)$ ou $v = (0, \beta)$, as quais são nulas. Se, porém, tomarmos $v = (\alpha, \beta)$, com $\alpha \neq 0$ e $\beta \neq 0$, veremos que não existe $(\partial f/\partial v)(0,0)$ pois

$$\frac{\partial f}{\partial v}(0,0) = \lim_{t \to 0} \frac{1}{t} \, \frac{t\alpha \cdot t\beta}{(t\alpha)^2 + (t\beta)^2} = \lim_{t \to 0} \frac{\alpha \cdot \beta}{t(\alpha^2 + \beta^2)}$$

e o último limite acima, evidentemente, não existe.

Ao contrário da maioria dos livros de Cálculo, em nossa definição de $\partial f/\partial v$ não supomos $|v| = 1$. Admitimos que $v \in \mathbb{R}^n$ seja um vetor arbitrário porque desejamos que $\partial f/\partial v$ dependa linearmente de v. Vejamos se isto realmente ocorre. Em primeiro lugar, se $0 \neq \alpha \in \mathbb{R}$ então existe $\partial f/\partial(\alpha v)$ num ponto $a \in U$ se, e somente se, existe $\partial f/\partial v$ e, no caso afirmativo, temos

$$\frac{\partial f}{\partial(\alpha v)}(a) = \lim_{t\to 0} \frac{f(a + t\alpha v) - f(a)}{t} = \alpha \cdot \lim_{t\to 0} \frac{f(a + t\alpha v) - f(a)}{t\alpha} = \\ = \alpha \cdot \frac{\partial f}{\partial v}(a).$$

Por outro lado, o exemplo abaixo mostra que a derivada direcional $\partial f/\partial v$ pode existir em todos os pontos do domínio de f, segundo todos os vetores $v \in \mathbb{R}^n$, sem que se tenha necessariamente

$$\frac{\partial f}{\partial(v + w)}(a) = \frac{\partial f}{\partial v}(a) + \frac{\partial f}{\partial w}(a).$$

Exemplo 3. Seja $g\colon \mathbb{R}^2 \to \mathbb{R}$ definida por $g(0,0) = 0$ e $g(x,y) = x^2y/(x^2 + y^2)$ quando $x^2 + y^2 \neq 0$. Um cálculo direto, a partir da definição, mostra que existem as derivadas direcionais $\partial g/\partial v$ em todos os pontos de $\mathbb{R}^2$, segundo qualquer vetor $v = (\alpha, \beta)$. Em particular, na origem, temos

$$\frac{\partial g}{\partial v}(0,0) = \lim_{t\to 0} \frac{g(t\alpha, t\beta)}{t} = \frac{\alpha^2\beta}{\alpha^2 + \beta^2}\cdot$$

Evidentemente, para $a = (0,0)$ não vale $\dfrac{\partial g}{\partial(v + w)}(a) = \dfrac{\partial g}{\partial v}(a) + \dfrac{\partial g}{\partial w}(a)$.

No §3, mostraremos que $\partial f/\partial v$ dependerá linearmente de v se f for "diferenciável", uma hipótese mais restritiva do que simplesmente possuir derivadas direcionais.

A função g do Exemplo 3 é contínua em todos os pontos do plano. [Isto é claro em $\mathbb{R}^2 - \{0\}$. Na origem, basta observar que $g(x,y) = x \cdot \cos\theta \cdot \operatorname{sen}\theta$. (Veja Exemplo 2.) Logo $\lim\limits_{x,y\to 0} g(x,y) = 0$.]

Não é verdade, porém, que a existência das derivadas direcionais implique em continuidade, como veremos agora.

Exemplo 4. Seja $h\colon \mathbb{R}^2 \to \mathbb{R}$ definida por $h(0,0) = 0$ e $h(x,y) = x^3y/(x^6 + y^2)$ se $(x,y) \neq (0,0)$. Em $\mathbb{R}^2 - \{0\}$, a função h é contínua.

Como, porém, $h(x, x^3) = 1/2$ para todo $x \neq 0$, vemos que h é descontínua na origem. Examinemos as derivadas direcionais. Para todo $v = (\alpha, \beta)$ temos:

$$\frac{\partial h}{\partial v}(0,0) = \lim_{t \to 0} \frac{h(t\alpha, t\beta)}{t} = \lim_{t \to 0} \frac{t^4\alpha^3\beta}{t^7\alpha^6 + t^3\beta^2} = \lim_{t \to 0} \frac{t\alpha^3\beta}{t^4\alpha^6 + \beta^2} = 0.$$

Assim, todas as derivadas direcionais $\partial h/\partial v$ na origem existem e dependem linearmente de v. O mesmo ocorre nos demais pontos $c \in \mathbb{R}^2 - \{0\}$, como se verifica mediante o cálculo elementar de $\dfrac{\partial h}{\partial v}(c) = \xi'(0)$, onde $\xi(t) = h(c + tv)$.

Outra propriedade desejável para um conceito adequado de derivada de uma função de n variáveis é que a composta de duas funções deriváveis seja ainda derivável. O próximo exemplo mostra uma função contínua $\varphi\colon \mathbb{R}^2 \to \mathbb{R}$, tal que $(\partial\varphi/\partial v)(z)$ existe para todo z e todo v, depende linearmente de v, mas $\varphi \circ \lambda$ não é diferenciável para um certo caminho diferenciável λ.

Exemplo 5. Definamos $\varphi\colon \mathbb{R}^2 \to \mathbb{R}$ pondo $\varphi(x,y) = x^3y/(x^4 + y^2)$ se $(x,y) \neq (0,0)$ e $\varphi(0,0) = 0$. Fora da origem, temos

$$\varphi(x,y) = x\frac{x^2}{\sqrt{x^4 + y^2}} \cdot \frac{y}{\sqrt{x^4 + y^2}} = x\frac{1}{\sqrt{1 + y^2/x^4}} \cdot \frac{\pm 1}{\sqrt{x^4/y^2 + 1}}\,,$$

logo $\lim\limits_{x,y \to 0} \varphi(x,y) = 0$. Assim, φ é contínua. Além disso, para todo $v = (\alpha, \beta)$, temos

$$\frac{\partial \varphi}{\partial v}(0,0) = \lim_{t \to 0} \frac{\varphi(t\alpha, t\beta)}{t} = \lim_{t \to 0} \frac{t\alpha^3\beta}{t^2\alpha^4 + \beta^2} = 0.$$

Portanto existem na origem, e dependem linearmente de v, todas as derivadas direcionais $\partial\varphi/\partial v$. Nos demais pontos $c \in \mathbb{R}^2 - \{0\}$ chega-se à mesma conclusão calculando-se $\dfrac{\partial\varphi}{\partial v}(c) = \xi'(0)$, com $\xi(t) = \varphi(c+tv)$, por meio das regras elementares de derivação. Entretanto, se considerarmos o caminho $\lambda\colon \mathbb{R} \to \mathbb{R}^2$, definido por $\lambda(t) = (t, t^2 \operatorname{sen} 1/t)$, $\lambda(0) = (0,0)$, veremos que λ é diferenciável e não existe a derivada $(\varphi \circ \lambda)'(0)$. Com efeito, seria

$$(\varphi \circ \lambda)'(0) = \lim_{t \to 0} \frac{\varphi(t, t^2 \operatorname{sen} 1/t)}{t} = \lim_{t \to 0} \frac{\operatorname{sen} 1/t}{1 + \operatorname{sen}^2 1/t}\,,$$

mas o último limite não existe.

A hipótese da existência de derivadas direcionais, embora fraca, não é inteiramente inócua. Ela permite demonstrar o Teorema do Valor Médio, o qual, para funções reais de n variáveis, é verdadeiro sob a forma de igualdade, como no caso de uma só variável.

Teorema do Valor Médio. *Seja* $f\colon U \to \mathbb{R}$ *definida no aberto* $U \subset \mathbb{R}^n$. *Suponhamos que o segmento de reta* $[a, a+v]$ *esteja contido em* U, *que a restrição* $f|[a, a+v]$ *seja contínua e que exista a derivada direcional* $\dfrac{\partial f}{\partial v}(x)$, *segundo* v, *em todo ponto* $x \in (a, a+v)$. *Então existe* $\theta \in (0,1)$ *tal que* $f(a+v) - f(a) = \dfrac{\partial f}{\partial v}(a+\theta v)$.

Demonstração: Definamos a função $\xi\colon [0,1] \to \mathbb{R}$ pondo $\xi(t) = f(a+tv)$. Pelas hipóteses feitas sobre f, ξ é contínua em $[0,1]$ e derivável em $(0,1)$. Pelo Teorema do Valor Médio para funções de uma variável real, existe $\theta \in (0,1)$ tal que $\xi(1) - \xi(0) = \xi'(\theta)$. Mas $\xi(1) = f(a+v)$, $\xi(0) = f(a)$ e

$$\begin{aligned}\xi'(\theta) &= \lim_{t\to 0} \frac{\xi(\theta+t) - \xi(\theta)}{t} = \lim_{t\to 0} \frac{f(a+(\theta+t)v) - f(a+\theta v)}{t} = \\ &= \lim_{t\to 0} \frac{f(a+\theta v+tv) - f(a+\theta v)}{t} = \frac{\partial f}{\partial v}(a+\theta v),\end{aligned}$$

o que demonstra o teorema.

Observação: A existência de $\partial f/\partial v$ em todo ponto de $(a, a+v)$ assegura apenas a continuidade de $f|(a, a+v)$ mas não de $f|[a, a+v]$.

Corolário. *Seja* $U \subset \mathbb{R}^n$ *aberto e conexo. Se* $f\colon U \to \mathbb{R}$ *possui derivadas direcionais em todo ponto* $x \in U$ *e* $\dfrac{\partial f}{\partial v}(x) = 0$ *para qualquer vetor* v, *então* f *é constante.*

Com efeito, fixemos $a \in U$. A existência de $\partial f/\partial v$ garante a continuidade da restrição $f|[a,b]$ para todo segmento de reta $[a,b]$ contido em U. Resulta então do Teorema do Valor Médio que $[a,b] \subset U$ implica $f(b) = f(a)$. Ora, qualquer ponto $x \in U$ pode (em virtude da conexidade do aberto U) ser ligado ao ponto a por uma poligonal contida em U, com vértices $a_0 = a, a_1, \ldots, a_k = x$. Temos sucessivamente $f(a) = f(a_1) = \cdots = f(x)$. Logo $f(x) = f(a)$ para todo $x \in U$, donde f é constante.

3 Funções diferenciáveis

A noção de função diferenciável, que apresentaremos agora, é devida a Fréchet e Stolz. Ela constitui, para funções de n variáveis, a extensão adequada do conceito de função derivável de uma só variável.

Dada $f\colon U \to \mathbb{R}$, com $U \subset \mathbb{R}^n$, seja $a \in U$. Diremos que a função f é *diferenciável* no ponto a quando existirem constantes $A_1, \ldots, A_n$ tais que, para todo vetor $v = (\alpha_1, \ldots, \alpha_n) \in \mathbb{R}^n$, com $a + v \in U$, se tenha

$$f(a+v) = f(a) + A_1 \cdot \alpha_1 + \cdots + A_n \cdot \alpha_n + r(v), \text{ onde } \lim_{v\to 0} \frac{r(v)}{|v|} = 0.$$

Quando f é diferenciável em todos os pontos de U, dizemos simplesmente que f é *diferenciável.*

Se f é diferenciável no ponto a então, tomando $v = te_i$, temos $\alpha_j = 0$ se $j \neq i$ e $\alpha_i = t$. Logo

$$\frac{f(a+te_i) - f(a)}{t} = A_i + \frac{r(te_i)}{t} = A_i \pm \frac{r(te_i)}{|te_i|}\cdot$$

Fazendo $t \to 0$, vemos que existe cada derivada parcial de f no ponto a, sendo $\frac{\partial f}{\partial x_i}(a) = A_i$. A definição abaixo é, portanto, equivalente à anterior.

Diremos que a função $f\colon U \to \mathbb{R}$ é diferenciável no ponto $a \in U$ quando existirem as derivadas parciais $\frac{\partial f}{\partial x_1}(a), \ldots, \frac{\partial f}{\partial x_n}(a)$ e, além disso, para todo vetor $v = (\alpha_1, \ldots, \alpha_n)$ tal que $a + v \in U$, tivermos

$$f(a+v) = f(a) + \frac{\partial f}{\partial x_i}(a) \cdot \alpha_1 + \cdots + \frac{\partial f}{\partial x_n}(a) \cdot \alpha_n + r(v),$$

onde $\lim_{v\to 0} \frac{r(v)}{|v|} = 0$.

Na igualdade acima, o “resto” $r(v)$ é definido como sendo igual a $f(a+v) - f(a) - \Sigma(\partial f/\partial x_i)(a) \cdot \alpha_i$. Esta definição pode ser dada para qualquer função que possua derivadas parciais. A essência da definição de diferenciabilidade é que, tomando $r(v)$ desta maneira, tem-se $\lim_{v\to 0} \frac{r(v)}{|v|} = 0$. Esta é a condição crucial, que deve ser verificada (direta ou indiretamente) sempre que quisermos provar que uma função é diferenciável.

De $\lim_{v\to 0}(r(v)/|v|) = 0$, concluimos que $\lim_{v\to 0} r(v) = 0$ pois $r(v) = (r(v)/|v|)|v|$. Daí resulta que toda função diferenciável num ponto é contínua nesse ponto. Com efeito, para $v = (\alpha_i, \ldots, \alpha_n)$, temos

$$\lim_{v\to 0}[f(a+v) - f(a)] = \lim_{v\to 0}\left[\Sigma\, \frac{\partial f}{\partial x_i}(a)\alpha_i + r(v)\right] = 0.$$

A condição $\lim_{v\to 0}(r(v)/|v|) = 0$ significa, entretanto, mais do que $r(v) \to 0$; ela quer dizer que $r(v)$ tende a zero mais rapidamente do que v, isto é, para valores de v suficientemente próximos de zero, a norma de $r(v)$ é uma fração arbitrariamente pequena da norma de v. Às vezes, isto se exprime dizendo-se que $r(v)$ é um infinitésimo de ordem superior a v. Assim, f é diferenciável no ponto a quando o acréscimo $f(a+v) - f(a)$ é igual a uma função linear de v, $\Sigma\, \frac{\partial f}{\partial x_i}(a)\cdot \alpha_i$, mais um resto infinitamente pequeno em relação a v.

Note que a validez da afirmação $\lim_{v\to 0}(r(v)/|v|) = 0$ independe da norma adotada em $\mathbb{R}^n$.

Em certas ocasiões, é preferível usar, em vez de $r(v)$, a função $\rho = \rho(v)$, definida para os valores de v tais que $a+v \in U$, do seguinte modo: $\rho(v) = r(v)/|v|$ se $v \neq 0$ e $\rho(0) = 0$. Então, a função f é diferenciável no ponto $a \in U$ se, e somente se, possui derivadas parciais nesse ponto e, para todo $v = (\alpha_1, \ldots, \alpha_n) \in R^n$ tal que $a+v \in U$, vale

$$f(a+v) = f(a) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)\cdot \alpha_i + \rho(v)\cdot |v|, \text{ onde } \lim_{v\to 0}\rho(v) = 0.$$

Assim, f é diferenciável no ponto a se, e somente se, a função real $\rho = \rho(v)$, definida pela igualdade acima (se $v \neq 0$) e por $\rho(0) = 0$, é contínua no ponto $v = 0$.

Para funções $f\colon I \to \mathbb{R}$, definidas num intervalo aberto $I \subset \mathbb{R}$, diferenciabilidade é o mesmo que derivabilidade, pois de $f(a+t) = f(a) + A\cdot t + \rho|t|$ se tira

$$\rho = \pm\left|\frac{f(a+t) - f(a)}{t} - A\right|,$$

logo $\lim_{t\to 0}\rho = 0$ se, e somente se, $A = f'(a)$.

Seja $f\colon U \to \mathbb{R}$ diferenciável no ponto $a \in U$. Já vimos que f é contínua e possui derivadas parciais nesse ponto. Mostraremos agora que

f admite derivada direcional segundo qualquer vetor $v = (\alpha_1, \ldots, \alpha_n)$, e vale a fórmula

$$\frac{\partial f}{\partial v}(a) = \frac{\partial f}{\partial x_1}(a) \cdot \alpha_1 + \cdots + \frac{\partial f}{\partial x_n}(a) \cdot \alpha_n\,.$$

Com efeito, para todo t suficientemente pequeno, temos $a + tv \in U$. Pela definição de diferenciabilidade, (as derivadas parciais sendo consideradas no ponto a) temos:

$$f(a+tv) = f(a) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \cdot t\alpha_i + \rho(tv) \cdot |t| \cdot |v|,$$

donde

$$\frac{f(a+tv) - f(a)}{t} = \Sigma\, \frac{\partial f}{\partial x_i} \cdot \alpha_i \pm \rho(tv) \cdot |v|.$$

Como $\lim\limits_{t \to 0} \rho(tv) = 0$, segue-se a fórmula enunciada.

Resulta da expressão $\partial f/\partial v = \Sigma(\partial f/\partial x_i)\alpha_i$ que se f é diferenciável num ponto então a derivada direcional $\partial f/\partial v$, nesse ponto, depende linearmente de v, isto é, não somente se tem $\partial f/\partial(\alpha v) = \alpha(\partial f/\partial v)$ como também $\partial f/\partial(v+w) = \partial f/\partial v + \partial f/\partial w$.

Uma propriedade relevante das funções diferenciáveis é dada pela

Regra da Cadeia. *Sejam $U \subset \mathbb{R}^m$, $V \subset \mathbb{R}^n$ abertos, $f = (f_1, \ldots, f_n)\colon U \to \mathbb{R}^n$ tal que $f(U) \subset V$ e cada função coordenada $f_k\colon U \to \mathbb{R}$ é diferenciável no ponto $a \in U$. Seja ainda $g\colon V \to \mathbb{R}$ uma função diferenciável no ponto $b = f(a)$. Então a função composta $g \circ f\colon U \to \mathbb{R}$ é diferenciável no ponto a e suas derivadas parciais são*

$$\frac{\partial(g \circ f)}{\partial x_i}(a) = \sum_{k=1}^{n} \frac{\partial g}{\partial y_k}(b) \cdot \frac{\partial f_k}{\partial x_i}(a).$$

Demonstração: Seja U_0 o conjunto dos vetores $v = (\alpha_1, \ldots, \alpha_m) \in \mathbb{R}^m$ tais que $a + v \in U$. Para $v \in U_0$ e $k = 1, \ldots, n$, temos

$$(1) \qquad f_k(a+v) = f_k(a) + \sum_{i=1}^{m} \frac{\partial f_k}{\partial x_i} \cdot \alpha_i + \rho_k \cdot |v|.$$

onde cada $\rho_k = \rho_k(v)$ é uma função definida em U_0, contínua no ponto 0, que se anula quando $v = 0$. [Acima, e no que se segue, as deri-

vadas $\partial f_k/\partial x_i$ e $\partial g/\partial y_k$ são consideradas nos pontos a e b respectivamente.] Consideremos a aplicação $w = (\beta_1, \ldots, \beta_n)\colon U_0 \to \mathbb{R}^n$, contínua no ponto 0, cujas funções-coordenada são definidas por

$$\beta_k(v) = \sum_{i=1}^{m} \frac{\partial f_k}{\partial x_i} \cdot \alpha_i + \rho_k \cdot |v|. \tag{2}$$

Adotando, por exemplo, a norma da soma, temos $|\alpha_i|/|v| \leq 1$ se $v \neq 0$, logo cada $|\beta_k|/|v|$, e portanto a função $|w|/|v|$, é limitada numa vizinhança do ponto $v = 0$. Escrevendo gf em vez de $g \circ f$, podemos afirmar, em virtude de (1), (2) e da diferenciabilidade de g no ponto $b = f(a)$, que, para todo $v \in U_0$, vale:

$$gf(a+v) = g(b+w) = g(b) + \sum_{k=1}^{n} \frac{\partial g}{\partial y_k} \cdot \beta_k + \sigma \cdot |w|,$$

onde $\sigma = \sigma(v)$ é uma função real contínua no ponto 0, que se anula no ponto $v = 0$ (pois w também se anula nesse ponto). Usando a definição de β_k, obtemos

$$\begin{aligned} gf(a+v) &= gf(a) + \sum_{k=1}^{n} \frac{\partial g}{\partial y_k} \left[\sum_{i=1}^{m} \frac{\partial f_k}{\partial x_i} \cdot \alpha_i + \rho_k |v| \right] + \sigma \cdot |w| = \\ &= gf(a) + \sum_{i=1}^{m} A_i \cdot \alpha_i + R, \end{aligned}$$

onde $A_i = \sum\limits_{k=1}^{n} \dfrac{\partial g}{\partial y_k} \cdot \dfrac{\partial f_k}{\partial x_i}$ e $R = \sum\limits_{i=1}^{n} \dfrac{\partial g}{\partial y_k} \cdot \rho_k \cdot |v| + \sigma |w|$.
Daí

$$\frac{R}{|v|} = \sum_{k=1}^{m} \frac{\partial g}{\partial y_k} \rho_k + \sigma \cdot \frac{|w|}{|v|}.$$

Quando v tende a zero, sabemos que cada função ρ_k tende a zero, que o quociente $|w|/|v|$ é limitado e que $\lim\limits_{v \to 0} \sigma = 0$. Segue-se que $\lim\limits_{v \to 0}(R/|v|) = 0$. Isto mostra que $g \circ f$ é diferenciável no ponto a e suas derivadas parciais são os números A_i.

Observação: A notação clássica do Cálculo Diferencial, às vezes imprecisa porém bastante sugestiva, além de compatível com a prática (então universal) de enfatizar *grandezas* ("y é uma função de x") em vez de

aplicações ("f leva x em y"), seria a seguinte, para a Regra da Cadeia: os pontos de U seriam escritos como "x"e os de V como "y"; as funções f_k seriam escritas como $y_k\,[= y_k(x)]$. A derivada $\partial(g \circ f)/\partial x_i$ seria a "derivada de g em relação à variável x_i", indicada com $\partial g/\partial x_i$. A Regra da Cadeia seria então:

$$\frac{\partial g}{\partial x_i} = \sum_k \frac{\partial g}{\partial y_k} \cdot \frac{\partial y_k}{\partial x_i}\,.$$

Não se pode negar a elegância nostálgica desta fórmula. Ela, porém, se acha demasiadamente comprometida com sistemas de coordenadas, para o gosto atual. No Capítulo 5, apresentaremos a versão intrínseca da Regra da Cadeia, cujo significado independe de coordenadas.

Corolário 1. *Se $f\colon U \to \mathbb{R}$ é diferenciável no ponto b e se $\lambda\colon (a-\varepsilon, a+\varepsilon) \to U \subset \mathbb{R}^n$ é um caminho diferenciável no ponto a, com $\lambda(a) = b$ e $\lambda(t) = (\lambda_1(t), \ldots, \lambda_n(t))$, então a função composta $f \circ \lambda\colon (a-\varepsilon, a+\varepsilon) \to \mathbb{R}$ é diferenciável no ponto a, e tem-se*

$$(f \circ \lambda)'(a) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(b) \cdot \lambda_i'(a).$$

Se escrevermos $\lambda(t) = (x_1(t), \ldots, x_n(t))$ então

$$\lambda'(t) = \left(\frac{dx_1}{dt}, \cdots, \frac{dx_n}{dt}\right).$$

Indicando com $\dfrac{df}{dt}$ a derivada da função composta

$$t \mapsto f(\lambda(t)) = f(x_1(t), \ldots, x_n(t)),$$

a Regra da Cadeia assume a forma clássica

$$\frac{df}{dt} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \cdot \frac{dx_i}{dt}\,.$$

Corolário 2. *Sejam $U \subset \mathbb{R}^n$ um conjunto aberto, $f\colon U \to \mathbb{R}$ diferenciável no ponto a, com $f(U) \subset I$, $g\colon I \to \mathbb{R}$ diferenciável no ponto $b = f(a)$. Então $g \circ f\colon U \to \mathbb{R}$ é diferenciável no ponto a e, para cada $i = 1, \ldots, n$ vale*

$$\frac{\partial(g \circ f)}{dx_i}(a) = g'(b) \cdot \frac{\partial f}{\partial x_i}(a).$$

Decorre da Regra da Cadeia que, se $f\colon U \to \mathbb{R}$ é diferenciável no ponto $a \in U$, ao calcularmos a derivada direcional $\dfrac{\partial f}{\partial v}(a) = (f \circ \lambda)'(0)$, não é necessário tomar $\lambda(t) = a + tv$. Em vez de nos restringirmos a um caminho retilíneo, podemos considerar *qualquer* caminho $\lambda\colon (-\varepsilon, \varepsilon) \to U$, *diferenciável no ponto* a, *com* $\lambda(0) = a$ *e* $\lambda'(0) = v = (\alpha_1, \ldots, \alpha_n)$ *e teremos ainda*

$$\frac{\partial f}{\partial v}(a) = (f \circ \lambda)'(0) = \lim_{t \to 0} \frac{f(\lambda(t)) - f(a)}{t} \cdot$$

Com efeito, pela Regra da Cadeia,

$$(f \circ \lambda)'(0) = \Sigma \, \frac{\partial f}{\partial x_i}(a) \cdot \lambda_i'(0) = \Sigma \, \frac{\partial f}{\partial x_i}(a) \cdot \alpha_i = \frac{\partial f}{\partial v}(a).$$

Revejamos agora os Exemplos 2, 3, 4 e 5 à luz da definição de diferenciabilidade. As funções que examinamos foram as seguintes:

$$f\colon \mathbb{R}^2 \to \mathbb{R}, \;\; f(x,y) = xy/(x^2 + y^2), \quad f(0,0) = 0;$$
$$g\colon \mathbb{R}^2 \to \mathbb{R}, \;\; g(x,y) = x^2y/(x^2 + y^2), \;\; g(0,0) = 0;$$
$$h\colon \mathbb{R}^2 \to \mathbb{R}, \;\; h(x,y) = x^3y/(x^6 + y^2), \;\; h(0,0) = 0;$$
$$\varphi\colon \mathbb{R}^2 \to \mathbb{R}, \;\; \varphi(x,y) = x^3y/(x^4 + y^2), \;\; \varphi(0,0) = 0.$$

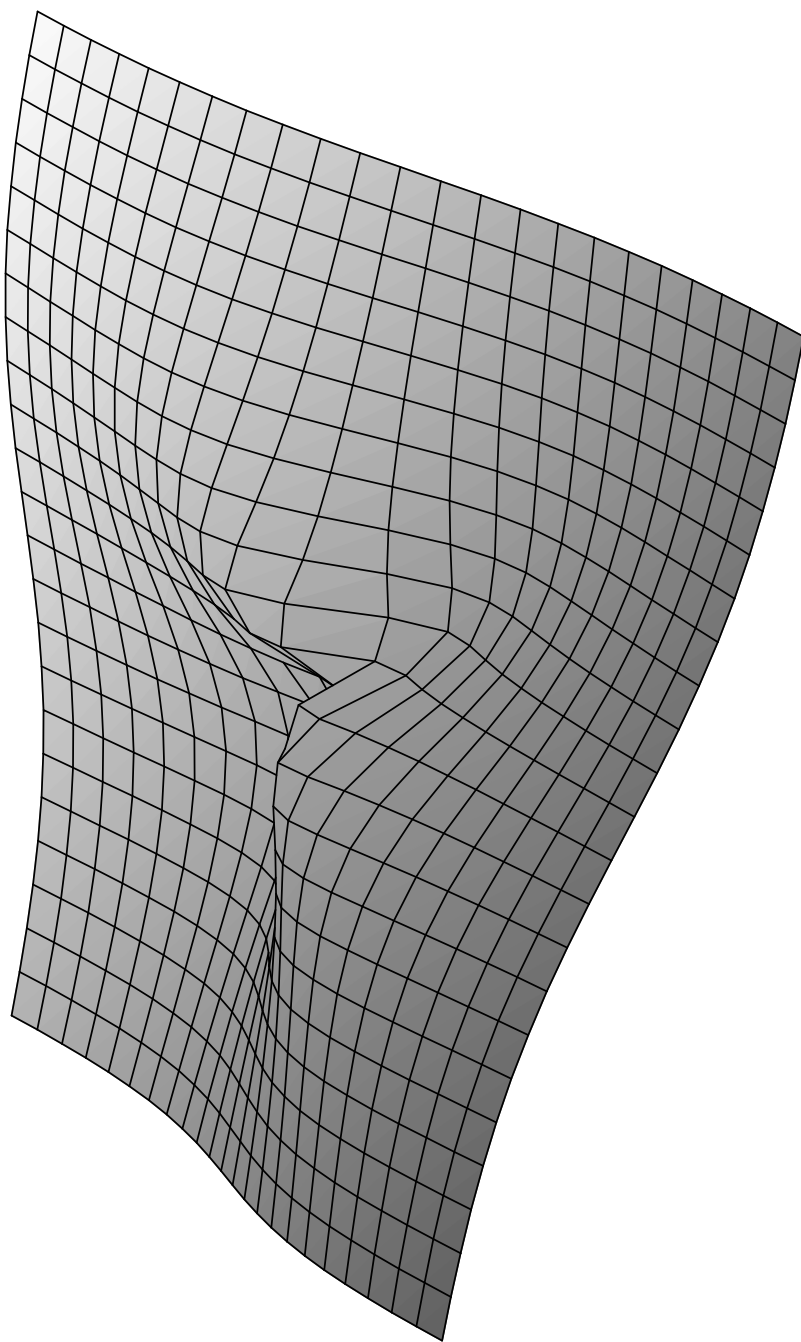
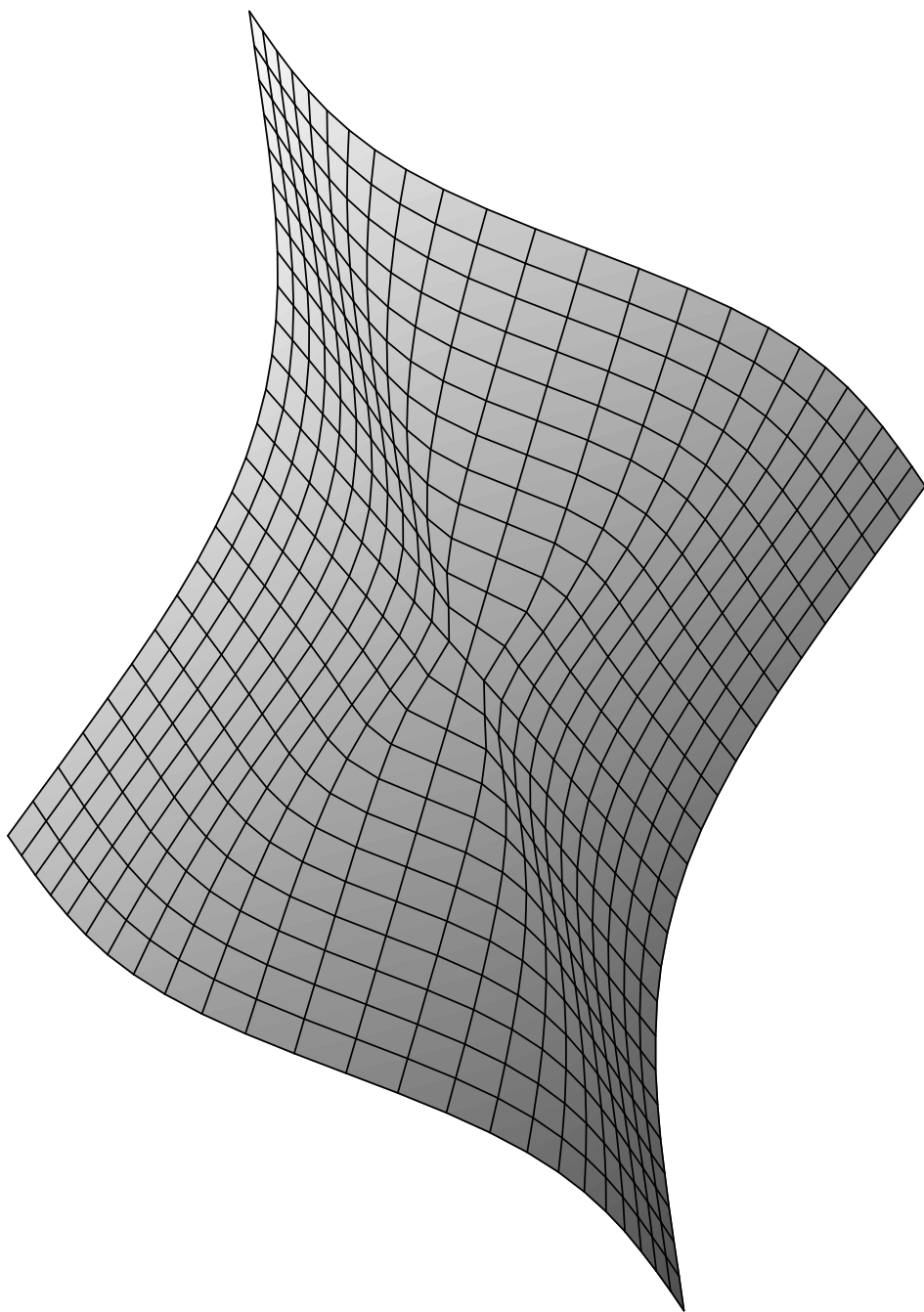



Gráfico de f.

Gráfico de g.

Nenhuma dessas funções é diferenciável na origem de $\mathbb{R}^2$: f porque não é contínua nem possui derivada direcional segundo todo vetor; g porque, embora sendo contínua e existindo $\partial g/\partial v(0,0)$ segundo todo vetor $v \in \mathbb{R}^2$, esta derivada não depende linearmente de v; h possui derivadas direcionais $\partial h/\partial v$, que dependem linearmente de v, mas não é contínua na origem; finalmente, φ é contínua em todo o plano, admite em todos os pontos do plano derivadas direcionais $\partial\varphi/\partial v$ que dependem linearmente de v, mas contraria a Regra da Cadeia porque, considerando o caminho $\lambda(t) = (t, t^2 \operatorname{sen} 1/t)$, *a* composta $\varphi \circ \lambda : \mathbb{R} \to \mathbb{R}$ não é derivável no ponto $t = 0$.

Estas são razões indiretas pelas quais as quatro funções acima não são diferenciáveis. A razão real é que, embora cada uma delas possua derivadas parciais na origem, elas não cumprem a condição

$$\lim_{v \to 0} \frac{r(v)}{|v|} = \lim_{\alpha,\beta \to 0} \frac{1}{\sqrt{\alpha^2 + \beta^2}} \left[f(\alpha, \beta) - \frac{\partial f}{\partial x} \cdot \alpha - \frac{\partial f}{\partial y} \cdot \beta \right] = 0,$$

onde as derivadas parciais são tomadas no ponto $(0,0)$. Acima, $\sqrt{\alpha^2 + \beta^2} = |v|$ é a norma euclidiana do vetor $v = (\alpha, \beta)$.

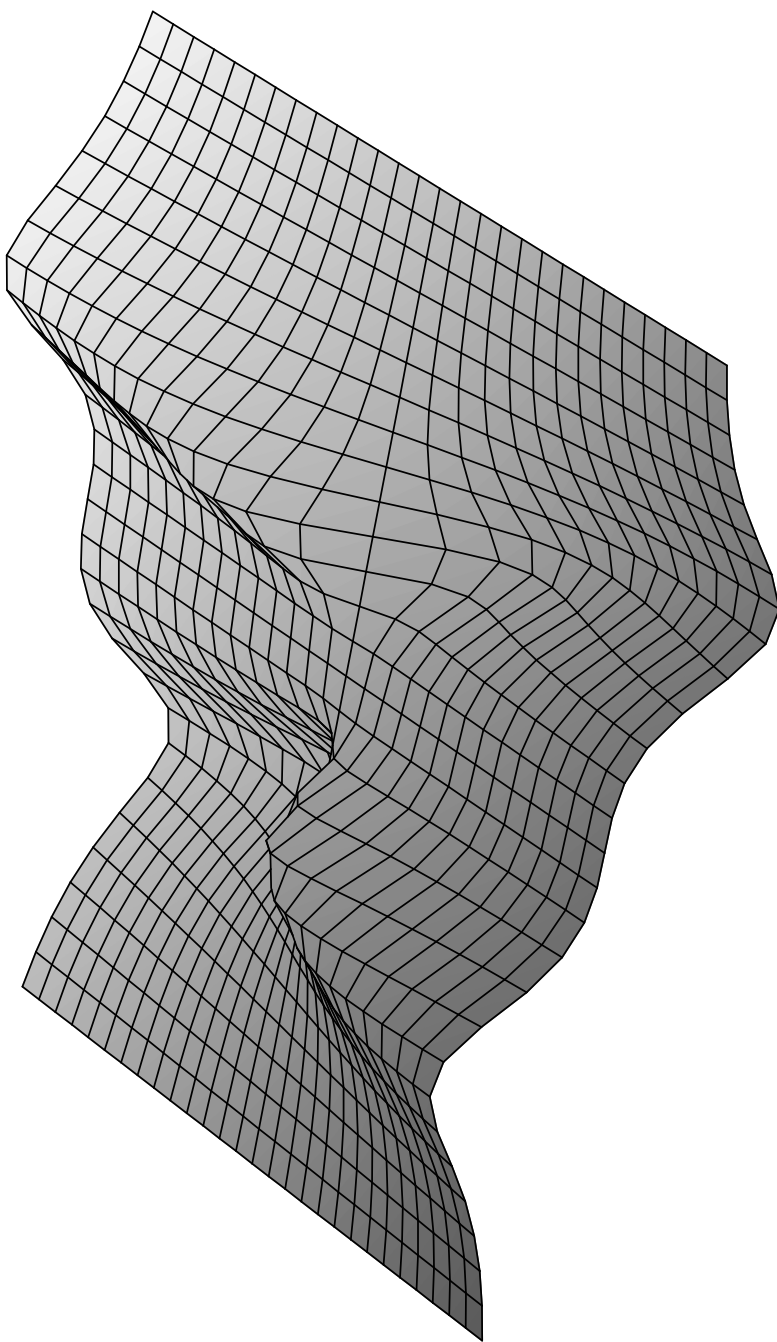



Gráfico de h.

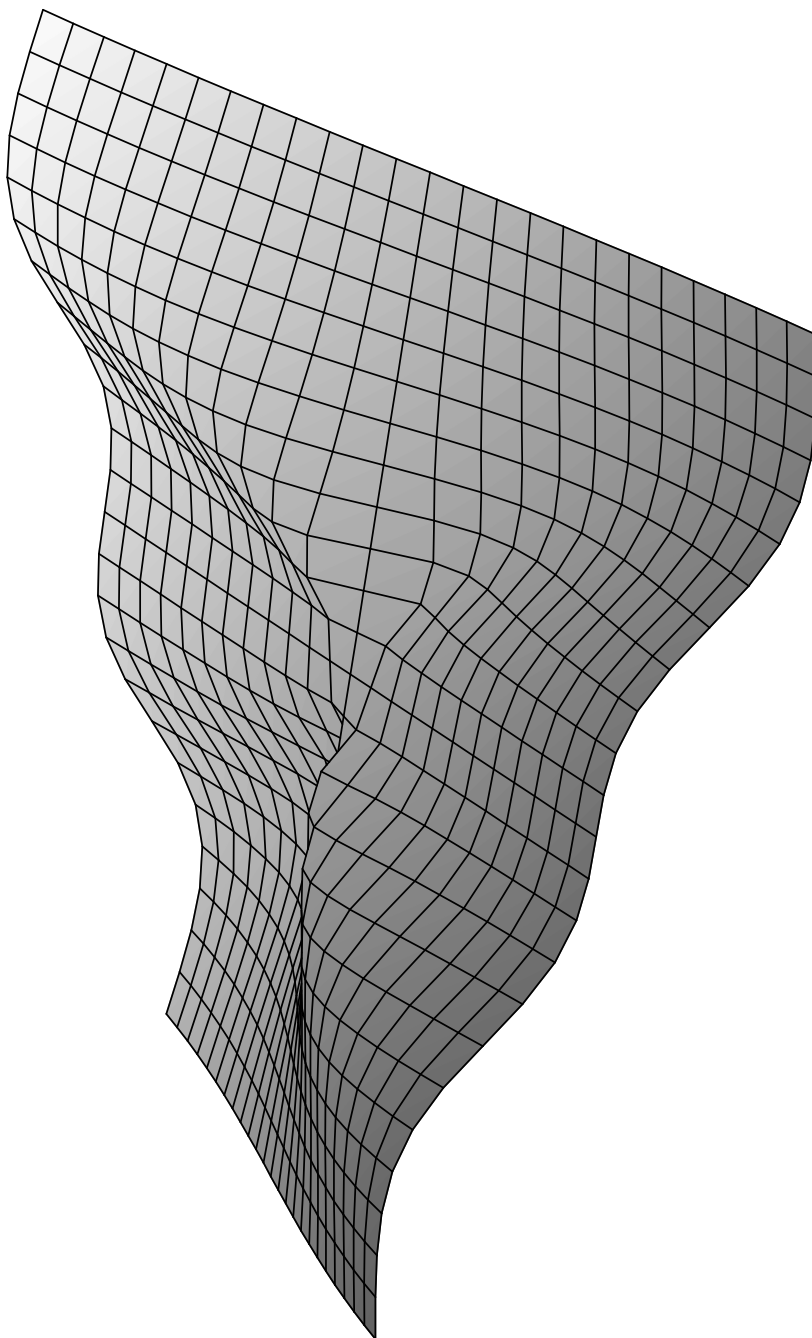



Gráfico de φ.

Exemplo 6. Uma função complexa $f\colon U \to \mathbb{C}$, definida num aberto $U \subset \mathbb{C}$, diz-se *derivável* no ponto $z = x + iy \in U$ quando existe o limite

$$\lim_{H\to 0} \frac{f(z+H) - f(z)}{H} = A,$$

o quociente acima sendo tomado no sentido dos números complexos. O número complexo $A = f'(z)$ chama-se a *derivada* da função complexa f no ponto z. A derivabilidade de f no ponto $z = x + iy$ equivale a dizer que

$$f(z+H) = f(z) + A \cdot H + r(H), \text{ onde } \lim_{H\to 0} \frac{r(H)}{|H|} = 0.$$

Sejam $A = a + ib$, $H = h + ik$ e $r = r_1 + ir_2$. Então f é derivável no ponto $z = x + iy$ se, e somente se,

$$(*) \quad f(z+H) = f(z) + (ah - bk) + i(bh + ak) + r_1(H) + i \cdot r_2(H),$$

onde $\lim\limits_{H\to 0} \dfrac{r_1(H)}{|H|} = \lim\limits_{H\to 0} \dfrac{r_2(H)}{|H|} = 0$.

Sejam $u, v\colon U \to \mathbb{R}$ as partes real e imaginária de f, ou seja $f(z) = u(z) + iv(z)$. Separando a parte real e a parte imaginária na igualdade (*) acima, obtemos

$$u(x+h, y+k) = u(x,y) + ah - bk + r_1(h,k),$$
$$\text{onde } \lim_{h,k\to 0} \frac{r_1(h,k)}{\sqrt{h^2+k^2}} = 0;$$

$$v(x+h, y+k) = v(x,y) + bh + ak + r_2(h,k),$$
$$\text{onde } \lim_{h,k\to 0} \frac{r_2(h,k)}{\sqrt{h^2+k^2}} = 0.$$

Assim, se a função complexa $f = u+iv$ é derivável no ponto $z = x+iv$ então sua parte real u e sua parte imaginária v são diferenciáveis no ponto (x,y) e, além disso, cumprem as condições $\dfrac{\partial u}{\partial x} = \dfrac{\partial v}{\partial y} (= a)$, $\dfrac{\partial u}{\partial y} = -\dfrac{\partial v}{\partial x} (= -b)$ nesse ponto. Estas são as chamadas *equações de Cauchy-Riemann.*

Reciprocamente, se $u, v: U \to \mathbb{R}$ são funções diferenciáveis no ponto $z = (x, y)$ e satisfazem nesse ponto as equações de Cauchy-Riemann $\partial u/\partial x = \partial v/\partial y$ e $\partial u/\partial y = -\partial v/\partial x$, então podemos reverter cada passo do argumento anterior e concluir que a função complexa $f = u + iv$ possui, no ponto $z = x + iy$, uma derivada complexa $f'(z)$, com $f'(z) = \dfrac{\partial u}{\partial x} - i\,\dfrac{\partial u}{\partial y} = \dfrac{\partial v}{\partial y} + i\,\dfrac{\partial v}{\partial x}\cdot$

A função complexa $f: U \to \mathbb{C}$ diz-se *holomorfa* quando possui derivada $f'(z)$ em todos os pontos do aberto U.

Uma função real $f: U \to \mathbb{R}$, definida no aberto $U \subset \mathbb{R}^n$, diz-se de *classe* C^1 quando existem, em cada ponto $x \in U$, as derivadas parciais $\dfrac{\partial f}{\partial x_1}(x), \ldots, \dfrac{\partial f}{\partial x_n}(x)$ e as n funções $\dfrac{\partial f}{\partial x_i}: U \to \mathbb{R}$, assim definidas, são contínuas. Mais geralmente, diremos que uma função $f: U \to \mathbb{R}$ é de *classe* C^k quando ela possuir derivadas parciais em todos os pontos de U e as funções $\dfrac{\partial f}{\partial x_1}, \ldots, \dfrac{\partial f}{\partial x_n}: U \to \mathbb{R}$ forem de classe C^{k-1}. Aqui, k é um inteiro > 0. Para completar a definição indutiva, diremos que uma função $f: U \to \mathbb{R}$ é de *classe* C^0 quando ela for contínua. Usaremos a notação $f \in C^k$. Escreveremos também $f \in C^\infty$, e diremos que f é de *classe* C^∞, quando $f \in C^k$ para todo $k \geq 0$. Evidentemente $C^0 \supset C^1 \supset \cdots \supset C^k \supset \cdots \supset C^\infty$, sendo todas estas inclusões estritas. (Veja Vol. 1, pág. 278.)

Teorema 1. *Se uma função $f: U \to \mathbb{R}$ possui derivadas parciais em todos os pontos do aberto $U \subset \mathbb{R}^n$ e cada uma delas é contínua no ponto c, então f é diferenciável no ponto c.*

Demonstração: Por simplicidade, consideraremos o caso $n = 2$. A situação geral se trata de modo análogo, apenas com notação mais complicada. Fixemos $c = (a, b) \in U$ e tomemos $v = (h, k)$ tal que $c + v \in U$. Seja

$$r(v) = r(h, k) = f(a + h,\, b + k) - f(a, b) - \frac{\partial f}{\partial x}\,h - \frac{\partial f}{\partial y}\,k,$$

onde as derivadas são calculadas no ponto $c = (a, b)$. Podemos escrever:

$$r(v) = f(a + h,\, b + k) - f(a, b + k) + f(a, b + k) - f(a, b) - \frac{\partial f}{\partial x}h - \frac{\partial f}{\partial y}k.$$

Pelo Teorema do Valor Médio para funções reais de uma variável real,

existem $\theta_1, \theta_2 \in [0,1]$ tais que

$$r(v) = \frac{\partial f}{\partial x}(a + \theta_1 h, b + k) \cdot h + \frac{\partial f}{\partial y}(a, b + \theta_2 k) \cdot k - \frac{\partial f}{\partial x} \cdot h - \frac{\partial f}{\partial x} \cdot k.$$

Logo

$$\begin{aligned} \frac{r(v)}{|v|} = & \left[\frac{\partial f}{\partial x}(a + \theta_1 \cdot h, b + k) - \frac{\partial f}{\partial x}(a, b)\right] \frac{h}{\sqrt{h^2 + k^2}} + \\ & + \left[\frac{\partial f}{\partial y}(a, b + \theta_2 \cdot k) - \frac{\partial f}{\partial y}(a, b)\right] \frac{k}{\sqrt{h^2 + k^2}} \cdot \end{aligned}$$

Ora, $h/\sqrt{h^2 + k^2}$ e $k/\sqrt{h^2 + k^2}$ estão, em valor absoluto, compreendidos entre 0 e 1. A continuidade das derivadas parciais nos dá então $\lim_{v \to 0} [r(v)/|v|] = 0$, logo f é diferenciável.

Corolário 1. *Toda função de classe C^1 é diferenciável.*

Observação: Escrevendo

$$r(v) = f(a+h, b+k) - f(a, b+k) - \frac{\partial f}{\partial x} \cdot h + f(a, b+k) - f(a, b) - \frac{\partial f}{\partial y} \cdot k,$$

vemos que existe $\theta \in (0,1)$ tal que

$$\begin{aligned} \frac{r(v)}{|v|} = & \left[\frac{\partial f}{\partial x}(a + \theta h, b + k) - \frac{\partial f}{\partial x}(a, b)\right] \frac{h}{|v|} + \\ & + \left[\frac{f(a, b + k) - f(a, b)}{k} - \frac{\partial f}{\partial y}(a, b)\right] \frac{k}{|v|} \cdot \end{aligned}$$

Como $h/|v|$ e $k/|v|$ têm valor absoluto menor do que um, a primeira parcela da soma acima tem limite zero quando $|v| \to 0$, desde que $\partial f/\partial x$ seja contínua no ponto (a, b). A segunda parcela também tem limite zero, em virtude da definição de derivada, sem que seja preciso supor $\partial f/\partial y$ contínua. Assim, para que uma função f, de duas variáveis, seja diferenciável num ponto, é suficiente que $\partial f/\partial x$ exista numa vizinhança do ponto e seja contínua no ponto, e que $\partial f/\partial y$ apenas exista no ponto em questão. Para funções de n variáveis, a diferenciabilidade num ponto é assegurada quando $n - 1$ das suas derivadas parciais são contínuas no ponto e a derivada parcial restante apenas existe ali.

Exemplo 7. Um polinômio em duas variáveis é uma função $p \colon \mathbb{R}^2 \to \mathbb{R}$, do tipo $p(x, y) = \Sigma a_{ij} x^i y^j$. Todo polinômio é evidentemente uma função

contínua e possui derivadas parciais $\partial p/\partial x = \Sigma i a_{ij} x^{i-1} y^j$, $\partial p/\partial y = \Sigma j a_{ij} x^i y^{j-1}$. Tais derivadas são ainda polinômios e portanto são funções contínuas em $\mathbb{R}^2$. Logo todo polinômio $p\colon \mathbb{R}^2 \to \mathbb{R}$ é uma função de classe C^1. Então as derivadas de p, sendo polinômios, são de classe C^1, portanto $p \in C^2$. Repetindo o mesmo argumento, concluimos que $p \in C^k$ para todo k, logo todo polinômio é, na realidade, uma função de classe C^∞. Afirmações semelhantes podem ser feitas sobre um polinômio a n variáveis, que é uma função $p\colon \mathbb{R}^n \to \mathbb{R}$, do tipo $p(x) = \Sigma\, a_{i_1 \ldots i_n} x_1^{i_1} \ldots x_n^{i_n}$.

A soma $f + g$, o produto $f \cdot g$ e o quociente f/g (se $g(x) \neq 0$ para todo x no domínio de g) de funções $f, g\colon U \to \mathbb{R}$, de classe C^k, são ainda funções de classe C^k.

Exemplo 8. Como exemplo de função diferenciável que não é de classe C^1, tomemos $f\colon \mathbb{R} \to \mathbb{R}$, com $f(x) = x^2 \cdot \operatorname{sen}(1/x)$ se $x \neq 0$, e $f(0) = 0$, lembrando que, para funções de uma variável, diferenciabilidade é o mesmo que existência da derivada.

Exemplo 9. O produto interno $f\colon \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$, $f(x, y) = \Sigma\, x_i y_i$, sendo um polinômo em $2m$ variáveis, é uma função de classe C^∞. Também é de classe C^∞ a função $g\colon \mathbb{R}^n \to \mathbb{R}$, $g(x) = |x|^2 = \Sigma\, x_i^2$, por ser ainda um polinômio, em n variáveis. Pela Regra da Cadeia, a norma euclidiana $h\colon \mathbb{R}^n \to \mathbb{R}$, $h(x) = |x| = \sqrt{\Sigma\, x_i^2}$, é de classe C^∞ quando restrita a $\mathbb{R}^n - \{0\}$. Na origem, a norma euclidiana não é diferenciável; nem sequer existem as derivadas parciais: $\dfrac{\partial h}{\partial x_i}(0^+) = 1$ e $\dfrac{\partial h}{\partial x_i}(0^-) = -1$. (Derivadas parciais laterais diferentes.) Quanto às normas que não provêm de um produto interno, elas podem não ser diferenciáveis mesmo em pontos $x \neq 0$. Por exemplo, seja $\xi\colon \mathbb{R}^2 \to \mathbb{R}$ a norma da soma, $\xi(x, y) = |x| + |y|$. Nos pontos $(x, 0)$ não existe $\dfrac{\partial \xi}{\partial y}$ e nos pontos $(0, y)$ não existe $\dfrac{\partial \xi}{\partial x}\cdot$

Para concluir, registremos um importante corolário do Teorema 1 (e da Regra da Cadeia), segundo o qual $g \circ f \in C^k$ desde que $g \in C^k$ e cada função coordenada de f também seja de classe C^k.

Corolário 2. *Sejam $U \subset \mathbb{R}^m$, $V \subset \mathbb{R}^n$ abertos, $f = (f_1, \ldots, f_n)\colon U \to \mathbb{R}^n$ tal que $f(U) \subset V$ e cada função coordenada $f_j\colon U \to \mathbb{R}$ é de classe C^k. Seja ainda $g\colon V \to \mathbb{R}$ uma função de classe C^k. Então a função composta $g \circ f\colon U \to \mathbb{R}$ é de classe C^k.*

Com efeito, pelo Corolário 1, g e cada f_j são diferenciáveis. (Estamos

supondo $k \geq 1$ pois o Corolário 2 é trivial se $k = 0$.) Podemos então aplicar a Regra da Cadeia, segundo a qual, para todo $i = 1, \ldots, m$, e todo $x \in U$:

$$\frac{\partial(g \circ f)}{\partial x_i}(x) = \sum_{j=1}^{n} \frac{\partial g}{\partial y_j}(f(x))\frac{\partial f_j}{\partial x_i}(x),$$

ou seja, vale a igualdade de funções

$$\frac{\partial(g \circ f)}{\partial x_i}(x) = \sum_{j=1}^{n} \left(\frac{\partial g}{\partial y_j} \circ f\right) \cdot \frac{\partial f_j}{\partial x_i} \,.$$

Suponhamos, por indução, que o Corolário 2 foi provado para classe C^{k-1}. Então, para cada $j = 1, \ldots, n$, a função composta $\dfrac{\partial g}{\partial y_j} \circ f$ é de classe C^{k-1}, o mesmo ocorrendo com $\dfrac{\partial f_j}{\partial x_i}$ para todo i, já que $f \in C^k$. Como o produto de funções de classe C^{k-1} é ainda desta classe, cada parcela da soma acima é de classe C^{k-1}, donde a soma também é. Assim, todas as derivadas parciais de $g \circ f$ são de classe C^{k-1}, portanto $g \circ f \in C^k$.

4 A diferencial de uma função

A derivada de um caminho $f : \mathbb{R} \to \mathbb{R}^m$ é um vetor. Na situação dual, o papel de derivada de uma função $f : \mathbb{R}^n \to \mathbb{R}$ é desempenhado por um funcional linear, conforme mostraremos agora.

Seja $f : U \to \mathbb{R}$ definida no aberto $U \subset \mathbb{R}^n$, diferenciável no ponto $a \in U$. A *diferencial* de f no ponto a é o funcional linear $df(a) : \mathbb{R}^n \to \mathbb{R}$, cujo valor no vetor $v = (\alpha_1, \ldots, \alpha_n)$ é dado por

$$df(a) \cdot v = \frac{\partial f}{\partial v}(a) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a) \cdot \alpha_i \,.$$

Como toda transformação linear $\mathbb{R}^n \to \mathbb{R}$, o funcional linear $df(a)$ possui uma matriz $1 \times n$ em relação à base canônica de $\mathbb{R}^n$. Se identificarmos o funcional com sua matriz, teremos

$$df(a) = \left(\frac{\partial f}{\partial x_1}(a), \ldots, \frac{\partial f}{\partial x_n}(a)\right).$$

Quando $f: U \to \mathbb{R}$ é diferenciável em todo ponto de U, obtemos uma aplicação $df: U \to (\mathbb{R}^n)^* = \mathcal{L}(\mathbb{R}^n; \mathbb{R})$, que associa a cada ponto $x \in U$ o funcional $df(x)$, cuja matriz é

$$\left(\frac{\partial f}{\partial x_1}(x), \ldots, \frac{\partial f}{\partial x_n}(x)\right).$$

A aplicação df é contínua se, e somente se, cada uma das suas funções coordenadas $\dfrac{\partial f}{\partial x_i}: U \to \mathbb{R}$ é contínua, isto é, se, e somente se, f é de classe C^1.

É comum indicar-se, em Análise, a base canônica de $(\mathbb{R}^n)^*$, com $(dx_1, \ldots, dx_n)$, logo $dx_i \cdot v = \alpha_i$ se $v = (\alpha_1, \ldots, \alpha_n)$. O motivo desta notação é o seguinte: como a i-ésima projeção $\pi_i: \mathbb{R}^n \to \mathbb{R}$ assume, em cada ponto $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ o valor $\pi_i(x) = x_i$, escreve-se x_i em vez de π_i. Calculando, de modo óbvio, a diferencial da i-ésima projeção $x_i: \mathbb{R}^n \to \mathbb{R}$, obtemos $dx_i(a) \cdot v = \alpha_i$ em todo ponto $a \in \mathbb{R}^n$. Escrevendo $dx_i \cdot v$ em vez de α_i, a definição da diferencial fica

$$df(a) \cdot v = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a) \cdot dx_i \cdot v.$$

Como esta igualdade vale para cada $v \in \mathbb{R}^n$, temos

$$df(a) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a) dx_i.$$

Isto significa que o funcional linear $df(a)$ se exprime como combinação linear dos funcionais dx_i, sendo $(\partial f/\partial x_i)(a)$ os coeficientes da combinação. Finalmente, a igualdade acima valendo para todo ponto $a \in U$, podemos escrever

$$df = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}\, dx_i.$$

A expressão formal da regra da cadeia (no caso $\mathbb{R} \to \mathbb{R}^n \to \mathbb{R}$) diz que, se cada coordenada x_i é função de um parâmetro real t então podemos "dividir ambos os membros da igualdade acima por dt" e obter

$$\frac{df}{dt} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \cdot \frac{dx_i}{dt}.$$

Todo funcional linear $\varphi\colon \mathbb{R}^n \to \mathbb{R}$ é diferenciável e, para todo $x \in \mathbb{R}^n$, $d\varphi(x) = \varphi$ (isto é, $d\varphi(x) \cdot v = \varphi \cdot v$). Com efeito, temos $\varphi(x) = c_1x_1 + \cdots + c_nx_n$, logo $\dfrac{\partial\varphi}{\partial x_i} = c_i$. Portanto $d\varphi(x) \cdot v = \Sigma \dfrac{\partial\varphi}{\partial x_i} \cdot \alpha_i = \Sigma\, c_i\alpha_i = \varphi \cdot v$.

Teorema 2. *Sejam $f, g\colon U \to \mathbb{R}$ diferenciáveis no ponto $a \in U$. Então:*

1) *$f + g\colon U \to \mathbb{R}$ é diferenciável e $d(f + g) = df + dg$;*

2) *$f \cdot g\colon U \to \mathbb{R}$ é diferenciável e $d(f \cdot g) = f \cdot dg + g \cdot df$;*

3) *Se $g(x) \neq 0$ para todo $x \in U$, então $f/g\colon U \to \mathbb{R}$ é diferenciável e $d(f/g) = (g \cdot df - f \cdot dg)/g^2$.*

Demonstração: As funções $s, m\colon \mathbb{R}^2 \to \mathbb{R}$, $q\colon \mathbb{R} \times (\mathbb{R} - \{0\}) \to \mathbb{R}$, definidas por $s(x, y) = x + y$, $m(x, y) = x \cdot y$ e $q(x, y) = x/y$, são de classe C^∞, logo diferenciáveis. A aplicação $F\colon U \to \mathbb{R}^2$, definida por $F(x) = (f(x), g(x))$, tem coordenadas diferenciáveis. Como $f + g = s \circ F$, $f \cdot g = m \circ F$, e $f/g = q \circ F$, a Regra da Cadeia assegura a diferenciabilidade de $f + g$, $f \cdot g$ e f/g. Além disso, temos

$$\begin{aligned}
\frac{\partial}{\partial x_i}(f + g) &= \frac{\partial f}{\partial x_i} + \frac{\partial g}{\partial x_i}\,, \\
\frac{\partial}{\partial x_i}(f \cdot g) &= f \cdot \frac{\partial g}{\partial x_i} + g \cdot \frac{\partial f}{\partial x_i} \quad \text{e} \\
\frac{\partial}{\partial x_i}\left(\frac{f}{g}\right) &= \frac{g \cdot \partial f/\partial x_i - f \cdot \partial g/\partial x_i}{g^2}\,.
\end{aligned}$$

Daí resultam as fórmulas enunciadas para as diferenciais.

O Teorema do Valor Médio, demonstrado no §2 para funções que admitem derivadas direcionais ao longo de um segmento, assume, para funções diferenciáveis, a forma abaixo, que decorre da anterior (bem como seu primeiro corolário):

Teorema do Valor Médio. *Seja $f\colon U \to \mathbb{R}$ diferenciável em todos os pontos do segmento de reta aberto $(a, a+v)$ e seja contínua sua restrição ao segmento fechado $[a, a + v] \subset U \subset \mathbb{R}^n$. Existe $\theta \in (0, 1)$ tal que*

$$f(a + v) - f(a) = df(a + \theta v) \cdot v = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a + \theta v) \cdot \alpha_i\,,$$

onde $v = (\alpha_1, \ldots, \alpha_n)$.

Corolário 1. *Seja $U \subset \mathbb{R}^n$ aberto e conexo. Se $f\colon U \to \mathbb{R}$ é diferenciável e $df(x) = 0$ (isto é, $\frac{\partial f}{\partial x_1}(x) = \cdots = \frac{\partial f}{\partial x_n}(x) = 0$) para todo $x \in U$ então f é constante.*

Corolário 2. *Sejam $U \subset \mathbb{R}^n$ um aberto convexo e $f\colon U \to \mathbb{R}$ uma função diferenciável. Se $|df(x)| \le M$ para todo $x \in U$ então, para quaisquer $x, y \in U$, temos $|f(x) - f(y)| \le M \cdot |x - y|$.*

Ou seja, num aberto convexo, toda função que possui diferencial limitada é Lipschitziana. No Corolário 2, $|df(x)|$ é a norma do funcional $df(x)\colon \mathbb{R}^n \to \mathbb{R}$, isto é, o maior dos números $\left|\frac{\partial f}{\partial v}(x)\right|$, para todo $v \in \mathbb{R}^n$, $|v| = 1$. Se tomarmos em $\mathbb{R}^n$ a norma euclidiana, ou a norma da soma, ou a norma do máximo, então $|df(x)|$ assume, respectivamente, os valores

$$\sqrt{\sum_i \left(\frac{\partial f}{\partial x_i}(x)\right)^2}, \ \text{máx} \cdot \left|\frac{\partial f}{\partial x_i}(x)\right| \ \text{ou} \ \sum_i \left|\frac{\partial f}{\partial x_i}(x)\right|.$$

Exemplo 10. Quando V não é convexo, uma função $g\colon V \to \mathbb{R}$ pode ter diferencial limitada em V e não ser Lipschitziana. Por exemplo, sejam f e U como no Exemplo 1 e tomemos $V = \{z \in U; |z| < 2\}$, $g = f|V$. Então $|dg(z)| \le 4$ para todo $z \in V$ mas g não é Lipschitziana pois, para todo $\varepsilon > 0$ temos $|g(1,\varepsilon) - g(1,-\varepsilon)| = 1$ enquanto $|(1,\varepsilon) - (1,-\varepsilon)| = 2\varepsilon$.

Uma conseqüência do Corolário 2 acima é que se $f\colon U \to \mathbb{R}$ é diferenciável e suas derivadas parciais são limitadas no aberto *convexo* $U \subset \mathbb{R}^n$, então f é uniformemente contínua em U. Em particular, f é a restrição de uma função (uniformemente) contínua $g\colon \overline{U} \to \mathbb{R}$. (Veja o Corolário do Teorema 14, Capítulo 1.)

5 O gradiente de uma função diferenciável

O produto interno natural induz um isomorfismo entre $\mathbb{R}^n$ e seu dual $(\mathbb{R}^n)^*$. Tal isomorfismo faz corresponder a cada vetor $v \in \mathbb{R}^n$ o funcional $v^* \in (\mathbb{R}^n)^*$ com $v^*(x) = \langle v, x\rangle$ para todo $x \in \mathbb{R}^n$. Se $v = (\alpha_1, \ldots, \alpha_n)$ então $v^*(e_1) = \alpha_1, \ldots, v^*(e_n) = \alpha_n$, logo a matriz de v^* em relação à base canônica de $\mathbb{R}^n$ é $(\alpha_1, \ldots, \alpha_n)$. A existência deste isomorfismo é responsável pelo fato de que no Cálculo Vetorial clássico (e portanto na Geometria e na Física tradicionais) não ocorrem funcionais lineares: em

vez de um funcional, toma-se o produto interno $\langle v, x\rangle$ de um vetor fixo v por um vetor variável x. A própria expressão $\varphi(x) = c_1x_1 + \cdots + c_nx_n$, que dá o valor do funcional φ no vetor $x = (x_1, \ldots, x_n)$, já indica isso: $\varphi(x)$ é o produto interno de x pelo vetor $v = (c_1, \ldots, c_n)$, ou seja, $\varphi = v^*$.

Dada a função diferenciável $f : U \to \mathbb{R}$, definida no aberto $U \subset \mathbb{R}^n$, definiremos o *gradiente* de f no ponto $a \in U$ como o vetor $\operatorname{grad} f(a)$, que corresponde ao funcional $df(a)$ segundo o isomorfismo acima descrito. Isto significa, por definição, que:

$$\langle \operatorname{grad} f(a), v\rangle = \frac{\partial f}{\partial v}(a) = df(a) \cdot v = \Sigma\, \frac{\partial f}{\partial x_i}(a) \cdot \alpha_i$$

para todo $v = (\alpha_1, \ldots, \alpha_n)$.

Em particular $\langle \operatorname{grad} f(a), e_i\rangle = \dfrac{\partial f}{\partial x_i}(a)$, logo

$$\operatorname{grad} f(a) = \left(\frac{\partial f}{\partial x_1}(a), \ldots, \frac{\partial f}{\partial x_n}(a)\right).$$

Se usarmos apenas bases ortonormais em $\mathbb{R}^n$, as coordenadas do vetor $\operatorname{grad} f(a)$ em relação à base $(e_1, \ldots, e_n)$ são as mesmas que as do funcional $df(a)$ com respeito à base dual $(dx_1, \ldots, dx_n)$. Nestas condições, o gradiente se torna praticamente indistinguivel da diferencial. Noutros contextos (como em Geometria Riemanniana), as duas noções são bem distintas. Mesmo no espaço euclidiano, o gradiente, por ser um vetor, apresenta aspectos geométricos muito convenientes para dar informações a respeito do comportamento da função, como vemos a seguir.

Destacaremos as três propriedades mais importantes do gradiente de uma função diferenciável f. *Nesta discussão, fixaremos um ponto a e suporemos que* $\operatorname{grad} f(a) \neq 0$. Então:

1º) O gradiente aponta para uma direção segundo a qual a função f é crescente;

2º) Dentre todas as direções ao longo das quais a função f cresce, a direção do gradiente é a de crescimento mais rápido;

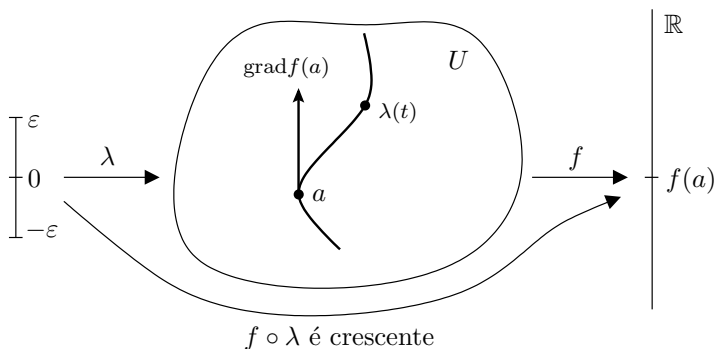
3º) O gradiente de f no ponto a é perpendicular à superfície de nível de f que passa por esse ponto.

Expliquemos essas afirmações.

Em primeiro lugar, se $w = \operatorname{grad} f(a)$ então

$$\frac{\partial f}{\partial w}(a) = \langle \operatorname{grad} f(a), w\rangle = |\operatorname{grad} f(a)|^2 > 0.$$

Isto significa que se $\lambda\colon (-\varepsilon,\varepsilon) \to U$ é um caminho diferenciável, com valores no domínio U da função f, tal que $\lambda(0) = a$ e $\lambda'(0) = \operatorname{grad} f(a)$, então a função real $t \mapsto f(\lambda(t))$ possui derivada positiva no ponto $t = 0$. Se supusermos f e λ de classe C^1, então a derivada de $f \circ \lambda$ será ainda positiva em todos os pontos de um intervalo aberto de centro 0, isto é, se tomarmos $\varepsilon > 0$ suficientemente pequeno então $f \circ \lambda\colon (-\varepsilon,\varepsilon) \to \mathbb{R}$ será uma função crescente. Isto é o que significa afirmar que "f cresce na direção do gradiente". Evidentemente, não se tem $\dfrac{\partial f}{\partial v}(a) > 0$ apenas quando $v = \operatorname{grad} f(a)$. Como $\dfrac{\partial f}{\partial v}(a) = \langle \operatorname{grad} f(a), v\rangle$, os vetores v que apontam para direções ao longo das quais a função cresce são aqueles que formam um ângulo agudo com $\operatorname{grad} f(a)$, isto é, tais que o produto interno $\langle \operatorname{grad} f(a), v\rangle$ é positivo. O que distingue o gradiente é o fato de que em sua direção o crescimento de f é mais rápido do que nas outras.

$f \circ \lambda$ é crescente

Isto quer dizer o seguinte: se v for um vetor tal que $|v| = |\operatorname{grad} f(a)|$ então

$$\frac{\partial f}{\partial v}(a) \le \frac{\partial f}{\partial(\operatorname{grad} f(a))}(a).$$

Com efeito, pela desigualdade de Schwarz:

$$\frac{\partial f}{\partial v}(a) = \langle \operatorname{grad} f(a), v\rangle \le |\operatorname{grad} f(a)|\,|v| =$$

$$= |\operatorname{grad} f(a)|^2 = \frac{\partial f}{\partial(\operatorname{grad} f(a))}(a).$$

Finalmente, esclareçamos a terceira das afirmações acima.

Dada $f: U \to \mathbb{R}$, diferenciável no aberto, $U \subset \mathbb{R}^n$, e dado um número real c, diz-se que o ponto $x \in U$ está *no nível* c, ou tem nível c, relativamente a f, quando $f(x) = c$. Fixado c, o conjunto dos pontos de U que estão no nível c é a imagem inversa $f^{-1}(c)$, a qual é chamada a *superfície de nível* c da função f. Quando $n = 2$, $f^{-1}(c)$ chama-se a *curva de nível* c de f. Convém, de início, chamar a atenção para o fato de que a imagem inversa $f^{-1}(c)$ às vezes não tem aspecto de curva ou superfície. (Por exemplo, $f: \mathbb{R}^2 \to \mathbb{R}$ pode ser constante, igual a c, num conjunto que contenha uma bola.) Melhor seria chamar $f^{-1}(c)$ de "conjunto de nível". Mas a terminologia está consagrada e se justifica devido a $f^{-1}(c)$ ser mesmo uma superfície (ou uma curva) sempre que $\operatorname{grad} f(x) \neq 0$ para todo x com $f(x) = c$, conforme provaremos adiante, com ajuda do teorema da função implícita.

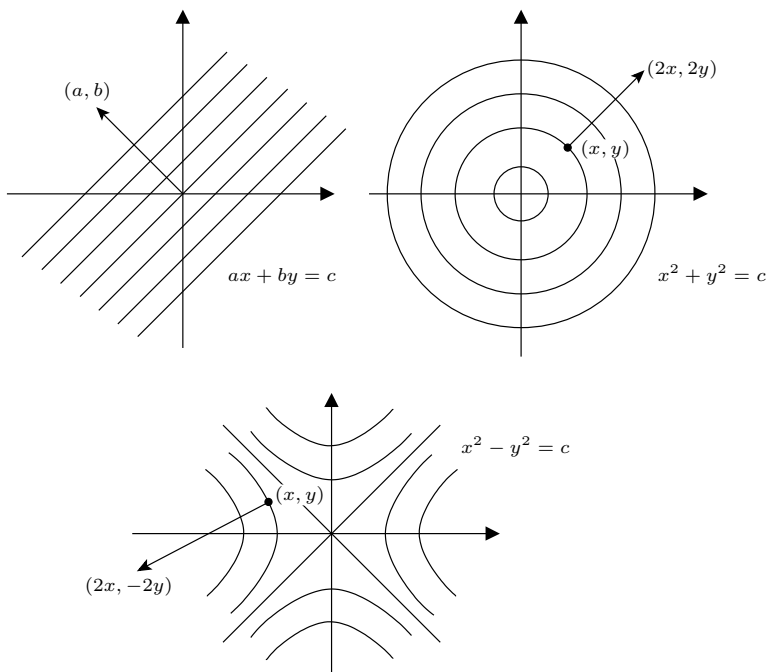
Dizer que um vetor w é perpendicular à superfície (ou curva) de nível $f^{-1}(c)$ no ponto a significa que w é perpendicular ao vetor velocidade, no ponto a, de qualquer caminho diferenciável no ponto $t = 0$, com $\lambda(0) = a$ e $\lambda(t) \in f^{-1}(c)$, isto é, $f(\lambda(t)) = c$, para todo $t \in (-\varepsilon, \varepsilon)$. Com efeito, desta última igualdade segue-se que

$$0 = (f \circ \lambda)'(0) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a) \cdot \lambda_i'(0) = \langle \operatorname{grad} f(a), \lambda'(0) \rangle,$$

logo $\operatorname{grad} f(a)$ é perpendicular a $\lambda'(0)$, vetor velocidade no ponto $a = \lambda(0)$ de qualquer caminho diferenciável λ, contido na superfície de nível de f que contém a.

Isto conlui a verificação das três propriedades do gradiente acima enunciadas. Vejamos agora alguns exemplos simples.

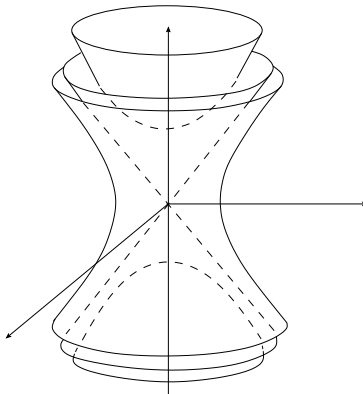
Exemplo 11. Sejam $f, g, h: \mathbb{R}^2 \to \mathbb{R}$ definidas por $f(x, y) = ax + by$ (onde $a^2 + b^2 \neq 0$), $g(x, y) = x^2 + y^2$ e $h(x, y) = x^2 - y^2$. As curvas de nível de f são a retas definidas pelas equações $ax + by = c$, para qualquer c real. O vetor gradiente de f é constante: $\operatorname{grad} f = (a, b)$ em qualquer ponto $(x, y) \in \mathbb{R}^2$. Assim, as curvas de nível de f são todas as retas perpendiculares ao vetor (a, b); tais retas são, evidentemente, paralelas umas às outras.

As curvas de nível c da função $g(x, y) = x^2 + y^2$ são as soluções de uma equação do tipo $g(x, y) = c$. Elas são vazias se $c < 0$. A curva de nível 0 reduz-se a um único ponto, a origem. Para $c > 0$, a curva de

nível c é o círculo de centro na origem e raio $\sqrt{c}$. O vetor gradiente de g no ponto (x, y) é $\operatorname{grad} g(x, y) = (2x, 2y)$, um vetor paralelo ao raio, o que é de se esperar, pois o raio é perpendicular a todo vetor tangente ao círculo naquele ponto.

A curva de nível c da função $h(x, y) = x^2 - y^2$ é formada por dois ramos de hipérbole quando $c \neq 0$. No caso de $c > 0$, a hipérbole $x^2 - y^2 = c$ tem como eixo o eixo das abcissas; quando $c < 0$ o eixo da hipérbole é o das ordenadas. Para $c = 0$, a curva de nível $x^2 - y^2 = 0$ consiste em duas retas perpendiculares que se cortam na origem: as diagonais do primeiro e terceiro e do segundo e quarto quadrantes, dadas por $y = x$ e $y = -x$, respectivamente. O gradiente da função h é o vetor $\operatorname{grad} h(x, y) = (2x, -2y)$. Atribuindo valores particulares a x e y, podemos observar que esse vetor é perpendicular à curva de nível que passa pelo ponto (x, y), e indica a direção de crescimento de h.

Note-se que, nos pontos onde o gradiente se anula (a origem na segunda e na terceira figura), ocorre uma quebra de regularidade na disposição das curvas de nível. Chamam-se *pontos singulares* ou *pontos*

críticos da função os pontos onde seu gradiente é o vetor zero.

Exemplo 12. Análogas tridimensionais das funções do exemplo anterior são $f, g, h\colon \mathbb{R}^3 \to \mathbb{R}$, definidas por $f(x, y, z) = ax + by + cz$ (com $a^2 + b^2 + c^2 \neq 0$), $g(x, y, z) = x^2 + y^2 + z^2$ e $h(x, y, z) = x^2 + y^2 - z^2$. As superfícies de nível de f são planos paralelos, todos perpendiculares ao

vetor (a, b, c), que é o gradiente de f. A superfície de nível c da função g é vazia se $c < 0$, reduz-se à origem se $c = 0$ e é uma esfera de centro na origem e raio $\sqrt{c}$ quando $c > 0$. A superfície de nível 0 da função h é um cone duplo, com vértice na origem e eixo no eixo dos z. Se $c > 0$, a superfície $x^2 + y^2 - z^2 = c$ é um hiperbolóide de revolução, com o mesmo eixo do cone. Se $c < 0$, $x^2 + y^2 - z^2 = c$ define um hiperbolóide de duas folhas.

6 A Regra de Leibniz

Regra de Leibniz (Derivação sob o sinal de integral.) *Dado $U \subset \mathbb{R}^n$, aberto, seja $f\colon U \times [a, b] \to \mathbb{R}$ uma função com as seguintes propriedades:*

1) *Para todo $x \in U$, a função $t \mapsto f(x, t)$ é integrável em $a \leq t \leq b$.*

2) *A i-ésima derivada parcial $\dfrac{\partial f}{\partial x_i}(x, t)$ existe para cada $(x, t) \in U \times [a, b]$ e a função $\dfrac{\partial f}{\partial x_i}\colon U \times [a, b] \to \mathbb{R}$, assim definida, é contínua.*

Então a função $\varphi\colon U \to \mathbb{R}$, *dada por* $\varphi(x) = \int_a^b f(x,t)dt$, *possui i-ésima derivada parcial em cada ponto* $x \in U$, *sendo*

$$\frac{\partial \varphi}{\partial x_i}(x) = \int_a^b \frac{\partial f}{\partial x_i}(x,t)\,dt.$$

Em suma: pode-se derivar sob o sinal de integral, desde que o integrando resultante seja uma função contínua.

Demonstração: Pelo Teorema do Valor Médio para funções reais de uma variável, se $x, x + se_i \in U$, existe $\theta \in (0,1)$ tal que

$$\begin{aligned}
&\frac{\varphi(x+se_i) - \varphi(x)}{s} - \int_a^b \frac{\partial f}{\partial x_i}(x,t)\,dt = \\
&= \int_a^b \left[\frac{f(x+se_i,t) - f(x,t)}{s} - \frac{\partial f}{\partial x_i}(x,t)\right] dt = \\
&= \int_a^b \left[\frac{\partial f}{\partial x_i}(x+\theta se_i,t) - \frac{\partial f}{\partial x_i}(x,t)\right] dt.
\end{aligned}$$

Pelo Teorema 21 b, Capítulo I, para todo $\varepsilon > 0$ dado, existe $\delta > 0$ tal que

$$|s| < \delta \Rightarrow \left|\frac{\partial f}{\partial x_i}(x+\theta se_i,t) - \frac{\partial f}{\partial x_i}(x,t)\right| < \varepsilon/(b-a),$$

seja qual for $t \in [a,b]$. Então $0 < |s| < \delta$ implica

$$\left|\frac{\varphi(x+sei) - \varphi(x)}{s} - \int_a^b \frac{\partial f}{\partial x_i}(x,t)\,dt\right| < \varepsilon.$$

como queríamos demonstrar.

Corolário. *Se* $f\colon U \times [a,b] \to \mathbb{R}$ *é contínua e possui n derivadas parciais contínuas* $\dfrac{\partial f}{\partial x_i}\colon U \times [a,b] \to \mathbb{R}$, *então* $\varphi\colon U \to \mathbb{R}$, *definida por* $\varphi(s) = \int_a^b f(x,t)\,dt$, *é de classe* C^1.

Com efeito, φ tem derivadas parciais contínuas. (Vide (12.7), no Capítulo I.)

Como conseqüência da Regra de Leibniz, prova-se o teorema da inversão da ordem nas integrais repetidas. Trata-se do seguinte: temos uma função contínua $f\colon [a,b] \times [c,d] \to \mathbb{R}$. Então, pela observação (12.7), logo após o Teorema 21 b, no Capítulo I, a função $\xi\colon [a,b] \to \mathbb{R}$,

definida por $\xi(s) = \int_c^d f(s,t)\,dt$, é contínua. A integral $\int_a^b \xi(s)\,ds$ se escreve como

$$\int_a^b \left[\int_c^d f(s,t)\,dt\right] ds \quad \text{ou} \quad \int_a^b ds \int_c^d f(s,t)\,dt,$$

tendo em mente a ordem em que foram efetuadas as integrações. Agora podemos enunciar o

Teorema da Inversão da Ordem nas Integrais Repetidas. *Para toda função contínua* $f\colon [a,b]\times[c,d]\to\mathbb{R}$, *vale*

$$\int_a^b ds \int_c^d f(s,t)\,dt = \int_c^d dt \int_a^b f(s,t)\,ds.$$

Demonstração: Definamos a função $\varphi\colon [a,b]\to\mathbb{R}$, pondo $\varphi(x) = \int_c^d dt \int_a^x f(s,t)\,ds$. Temos $\varphi(a) = 0$, $\varphi(b) = \int_c^d dt \int_a^b f(s,t)\,ds$ e, pela Regra de Leibniz $\varphi'(x) = \int_c^d f(x,t)\,dt$, pois o integrando $f(x,t)$ é uma função contínua de (x,t). Pelo Teorema Fundamental do Cálculo, segue-se que $\varphi(b) = \varphi(a) + \int_a^b \varphi'(s)\,ds = \int_a^b ds \int_c^d f(s,t)\,dt$, o que demonstra o teorema.

Outra conseqüência interessante da Regra de Leibniz é uma demonstração do Teorema de Schwarz, que apresentaremos no §7, adiante.

Exemplo 12a. Seja $f\colon U\times[a,b]\to\mathbb{R}$ contínua, com derivadas parciais contínuas $\dfrac{\partial f}{\partial x_1},\ldots,\dfrac{\partial f}{\partial x_n}\colon U\times[a,b]\to\mathbb{R}$. Seja $g\colon U\to[a,b]$ de classe C^1. Vamos mostrar que a função $\varphi\colon U\to\mathbb{R}$, definida por $\varphi(x) = \int_a^{g(x)} f(x,t)\,dt$, é de classe C^1, e suas derivadas parciais são expressas pela fórmula:

$$\frac{\partial\varphi}{\partial x_i}(x) = \int_a^{g(x)} \frac{\partial f}{\partial x_i}(x,t)\,dt + \frac{\partial g}{\partial x_i}(x)\cdot f(x,g(x)).$$

Com efeito, se consideramos a função $\xi\colon U\times[a,b]\to\mathbb{R}$, definida por $\xi(x,u) = \int_a^u f(x,t)\,dt$, a Regra de Leibniz e a regra de derivação de uma integral indefinida dão

$$\frac{\partial\xi}{\partial x_i}(x,u) = \int_a^u \frac{\partial f}{\partial x_i}(x,t)\,dt \quad \text{e} \quad \frac{\partial\xi}{\partial u}(x,u) = f(x,u).$$

Assim, ξ é de classe C^1, portanto diferenciável. Podemos então usar a Regra da Cadeia, segundo a qual a função composta $\varphi(x) = \xi(s, g(x))$ tem como derivada parcial

$$\frac{\partial \varphi}{\partial x_i}(x) = \frac{\partial \xi}{\partial x_i}(x, g(x)) + \frac{\partial g}{\partial x_i}(x) \cdot \frac{\partial \xi}{\partial u}(x, g(x)) =$$
$$= \int_a^{g(x)} \frac{\partial f}{\partial x_i}(x,t)\,dt + \frac{\partial g}{\partial x_i}(x) \cdot f(x, g(x)).$$

Daí decorrem fórmulas análogas para as derivadas parciais das funções $\psi(x) = \int_{g(x)}^{b} f(x,t)\,dt$ e $\zeta(x) = \int_{g(x)}^{h(x)} f(x,t)\,dt$.

Exemplo 13. Seja $f\colon I \to \mathbb{R}$ uma função contínua, definida num intervalo I que contém 0. Para cada inteiro $n \geq 1$, seja $F_n\colon I \to \mathbb{R}$ a função definida por

$$F_n(x) = \int_0^x \frac{(x-t)^{n-1}}{(n-1)!} f(t)\,dt.$$

Para completar, seja $F_0 = f$. A fórmula acima deduzida nos permite facilmente calcular a derivada de cada F_n. Obtemos $F_n' = F_{n-1}$. Portanto F_n é uma função que se anula para $x = 0$ e cuja n-ésima derivada é igual a f. (É interessante notar que F_n foi obtida com uma única integração.)

7 O Teorema de Schwarz

Seja $f\colon U \to \mathbb{R}$ diferenciável no aberto $U \subset \mathbb{R}^n$. Para $i = 1, \ldots, n$, cabe indagar se as funções $\dfrac{\partial f}{\partial x_i}\colon U \to \mathbb{R}$ são diferenciáveis num ponto $a \in U$. Se todas são, diz-se que f é *duas vezes diferenciável* no ponto a. Neste caso, para todos os inteiros $i, j = 1, 2, \ldots, n$, existem as *derivadas parciais de segunda ordem*

$$\frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right)(a) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a).$$

Quando f é duas vezes diferenciável em todos os pontos de U, ficam definidas n^2 funções $\dfrac{\partial^2 f}{\partial x_j \partial x_i}\colon U \to \mathbb{R}$, $\quad 1 \leq i,\, j \leq n$. Se todas essas funções são diferenciáveis num ponto, diz-se que f é *três vezes diferenciável* naquele ponto. E assim por diante.

Sabemos que $f \in C^1 \Rightarrow f$ diferenciável. Daí, e das definições que acabamos de dar, resulta que $f \in C^k \Rightarrow f$ é k vezes diferenciável.

O Teorema de Schwarz assegura que, mediante algumas hipóteses naturais, a ordem em que são tomadas as derivadas parciais repetidas não influi no resultado final. Há diferentes enunciados para esse teorema. Daremos dois. O primeiro é o mais simples e natural.

Teorema de Schwarz. *Seja $f\colon U \to \mathbb{R}$ duas vezes diferenciável no ponto $c \in U \subset \mathbb{R}^n$. Para quaisquer $1 \le i, j \le n$, tem-se*

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(c) = \frac{\partial^2 f}{\partial x_j \partial x_i}(c).$$

Demonstração: Para simplificar a notação, e sem perder a generalidade, podemos supor $U \subset \mathbb{R}^2$, $c = (a,b)$ e provar que $\frac{\partial^2 f}{\partial x \partial y}(a,b) = \frac{\partial^2 f}{\partial y \partial x}(a,b)$. Existe $\varepsilon > 0$ tal que o quadrado $(a-\varepsilon, a+\varepsilon) \times (b-\varepsilon, b+\varepsilon)$ está contido em U. Para todo $t \in (-\varepsilon, \varepsilon)$, ponhamos

$$\varphi(t) = f(a+t, b+t) - f(a+t, b) - f(a, b+t) + f(a,b).$$

Então, escrevendo $\xi(x) = f(x, b+t) - f(x,b)$, vemos que $\varphi(t) = \xi(a+t) - \xi(a)$. Pelo Teorema do Valor Médio para funções de uma variável, existe $\theta \in (0,1)$ tal que $\varphi(t) = \xi'(a+\theta t) \cdot t$, ou seja:

$$\varphi(t) = \left[\frac{\partial f}{\partial x}(a+\theta t, b+t) - \frac{\partial f}{\partial x}(a+\theta t, b)\right] t.$$

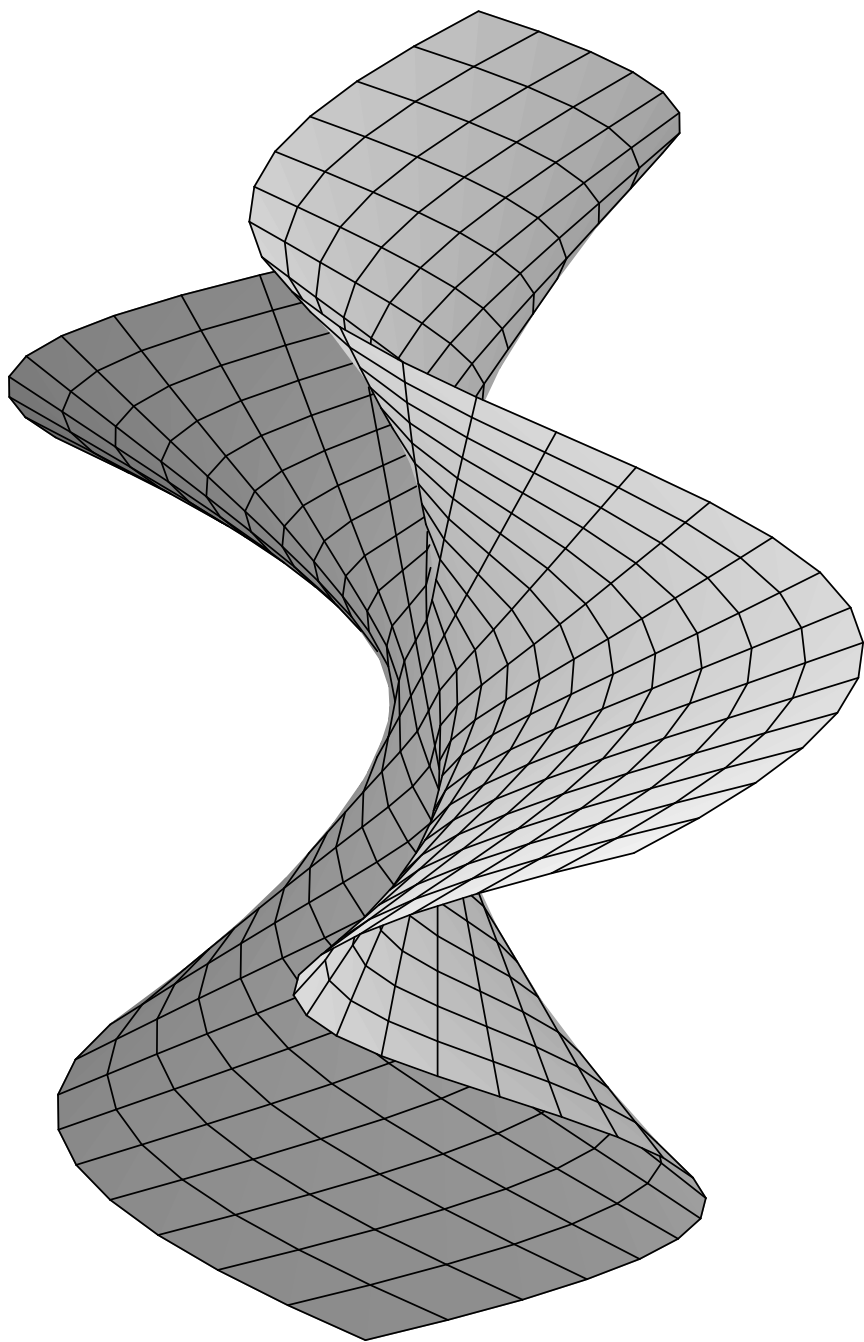
Como a função $\frac{\partial f}{\partial x}\colon U \to \mathbb{R}$ é diferenciável no ponto $c = (a,b)$, temos as igualdades: (As derivadas não explicitadas abaixo são tomadas no ponto c.)

$$\frac{\partial f}{\partial x}(a+\theta t, b+t) = \frac{\partial f}{\partial x} + \frac{\partial^2 f}{\partial x^2} \cdot \theta t + \frac{\partial^2 f}{\partial y \partial x} \cdot t + \rho_1 \cdot t, \text{ onde } \lim_{t \to 0} \rho_1 = 0;$$

$$\frac{\partial f}{\partial x}(a+\theta t, b) = \frac{\partial f}{\partial x} + \frac{\partial^2 f}{\partial x^2} \cdot \theta t + \rho_2 \cdot t, \text{ onde } \lim_{t \to 0} \rho_2 = 0.$$

Daí $\varphi(t) = \frac{\partial^2 f}{\partial y \partial x} \cdot t^2 + \rho \cdot t^2$, onde $\rho = \rho_1 - \rho_2$ e portanto $\lim_{v \to 0} \rho = 0$. Segue-se então que $\lim_{t \to 0} \frac{\varphi(t)}{t^2} = \frac{\partial^2 f}{\partial y \partial x}(a,b)$. Um raciocínio semelhante,

usando a função $\zeta(y) = f(a+t, y) - f(a, y)$, nos conduziria a $\lim_{t \to 0} \frac{\varphi(t)}{t^2} = \frac{\partial^2 f}{\partial x \partial y}(a, b)$. Isto prova o teorema.

Exemplo 14. O caso mais conhecido de uma função que possui derivadas de segunda ordem em todos os pontos do plano mas $\frac{\partial^2 f}{\partial x \partial y} \neq \frac{\partial^2 f}{\partial y \partial x}$ na origem é $f\colon \mathbb{R}^2 \to \mathbb{R}$, definida por $f(x, y) = \frac{xy(x^2 - y^2)}{x^2 + y^2}$ se $x^2+y^2 \neq 0$ e $f(0, 0) = 0$.

Nos pontos de $\mathbb{R}^2 - \{0\}$, f possui derivadas parciais repetidas de to-

das as ordens. Na origem, existem (e são nulas) $\partial^2 f/\partial x^2$ e $\partial^2 f/\partial y^2$. Quanto às derivadas mistas, começamos observando que, para todo $y \neq 0$, temos $f(0, y) = 0$, logo

$$\frac{\partial f}{\partial x}(0, y) = \lim_{x \to 0} \frac{xy(x^2 - y^2)}{x(x^2 + y^2)} = -y,$$

e daí

$$\frac{\partial^2 f}{\partial y \partial x}(0, 0) = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)(0, 0) = \lim_{y \to 0} \frac{\partial f/\partial x(0, y)}{y} = -1.$$

De modo análogo, vemos que $\dfrac{\partial^2 f}{\partial x \partial y}(0, 0) = 1$. Logo as duas derivadas mistas de f na origem são diferentes.

Como foi mencionado no §6, a seguinte versão do Teorema de Schwarz decorre da Regra de Leibniz.

Teorema 3. *Seja $f\colon U \to \mathbb{R}$ tal que existem $\dfrac{\partial f}{\partial x_i}$ e $\dfrac{\partial^2 f}{\partial x_i \partial x_j}$ em todos os pontos de U. Se as funções $\dfrac{\partial f}{\partial x_i}, \dfrac{\partial^2 f}{\partial x_i \partial x_j}\colon U \to \mathbb{R}$ são contínuas, então a derivada $\partial^2 f/\partial x_j \partial x_i$ existe em todos os pontos de U e vale $\partial^2 f/\partial x_i \partial x_j = \partial^2 f/\partial x_j \partial x_i$.*

Demonstração: Sem perda de generalidade, podemos supor que $U = I \times J$ é um retângulo em $\mathbb{R}^2$. Tomando um ponto $b \in J$, o Teorema Fundamental do Cálculo nos permite escrever, para todo ponto $(x, y) \in U$:

$$f(x, y) = f(x, b) + \int_b^y \frac{\partial f}{\partial y}(x, t)\, dt.$$

A continuidade de $\partial^2 f/\partial x \partial y$, admitida por hipótese, faz com que a Regra de Leibniz seja aplicável. Derivando em relação a x:

$$\frac{\partial f}{\partial x}(x, y) = \frac{\partial f}{\partial x}(x, b) + \int_b^y \frac{\partial^2 f}{\partial x \partial y}(x, t)\, dt.$$

Derivando em seguida relativamente a y, obtemos

$$\frac{\partial^2 f}{\partial y \partial x}(x, y) = \frac{\partial^2 f}{\partial x \partial y}(x, y),$$

pois $(\partial f/\partial x)(x, b)$ não depende de y e o integrando na segunda parcela é contínuo.

Observação: Não é muito surpreendente que, assim como a permutabilidade de integrações sucessivas, também uma versão do Teorema de Schwarz decorra da Regra de Leibniz. Afinal de contas, esses teoremas asseguram que, sob certas condições, é permitido inverter a ordem de dois limites sucessivos. Uma boa quantidade de resultados importantes em Análise decorrem da possibilidade de se permutarem limites, mediante hipóteses convenientes.

O caso mais simples do teorema acima ocorre quando $f \in C^2$. Então ele se torna um corolário da versão anterior, para funções duas vezes diferenciáveis.

O Teorema de Schwarz, evidentemente, se aplica para derivadas repetidas de ordem superior à segunda. Consideremos, por exemplo, uma função $f\colon U \to \mathbb{R}$, três vezes diferenciável no aberto $U \subset \mathbb{R}^2$. Há 6 derivadas mistas de terceira ordem para a função f, a saber

$$\frac{\partial^3 f}{\partial x \partial x \partial y}, \frac{\partial^3 f}{\partial x \partial y \partial x}, \frac{\partial^3 f}{\partial y \partial x \partial x},$$
$$\frac{\partial^3 f}{\partial y \partial y \partial x}, \frac{\partial^3 f}{\partial y \partial x \partial y}, \frac{\partial^3 f}{\partial x \partial y \partial y}.$$

Na realidade, porém, apenas duas delas podem ser diferentes, pois as 3 derivadas da primeira linha são iguais, bem como as da segunda linha. Com efeito, pelo Teorema de Schwarz:

$$\frac{\partial^3 f}{\partial x \partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial^2 f}{\partial x \partial y}\right) = \frac{\partial}{\partial x}\left(\frac{\partial^2 f}{\partial y \partial x}\right) = \frac{\partial^3 f}{\partial x \partial y \partial x}$$

e, por outro lado, escrevendo $g = \partial f/\partial x$, temos:

$$\frac{\partial^3 f}{\partial x \partial y \partial x} = \frac{\partial}{\partial x}\left[\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)\right] = \frac{\partial^2 g}{\partial x \partial y} = \frac{\partial^2 g}{\partial y \partial x} =$$
$$= \frac{\partial}{\partial y}\left[\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right)\right] = \frac{\partial^3 f}{\partial y \partial x \partial x}.$$

Analogamente para as 3 derivadas da segunda linha.

Por simplicidade, escreveremos $\partial^3 f/\partial x^2 \partial y$ para indicar qualquer das 3 derivadas da primeira linha e $\partial^3 f/\partial y^2 \partial x$ para indicar a outra derivada mista de 3ª ordem de f.

O processo se aplica em geral: se $f\colon U \to \mathbb{R}$ é uma função p-vezes diferenciável no aberto $U \subset \mathbb{R}^m$ então, para toda seqüência de inteiros não-negativos $i_1, \dots, i_n$, com $i_1 + \cdots + i_n = \alpha \le p$, a derivada de ordem α, $\partial^\alpha f/\partial x_1^{i_1}\, \partial x_2^{i_2} \dots \partial x_n^{i_n}$, que consiste em derivar i_1 vezes em relação a $x_1, \dots, i_n$ vezes em relação a x_n (podendo alguns dos i_k serem iguais a zero), não depende da ordem em que essas derivações foram efetuadas. Este é o enunciado geral do Teorema de Schwarz. Sua demonstração se baseia no caso que demonstramos e na observação de que qualquer mudança de ordem numa seqüência pode ser levada a efeito através de sucessivas transposições entre dois termos consecutivos da seqüência.

8 Fórmula de Taylor; pontos críticos

A fórmula de Taylor para uma função $f\colon U \to \mathbb{R}$, definida no aberto $U \subset \mathbb{R}^n$, é a seguinte:

$$f(a+v) = f(a) + df(a)\cdot v + \frac{1}{2}\, d^2 f(a)\cdot v^2 + \cdots + \frac{1}{p!}\, d^p f(a)\cdot v^p + r_p(v).$$

Conforme as hipóteses feitas e os objetivos desejados, temos, como no caso de funções de uma só variável, três situações principais:

1) **Fórmula de Taylor Infinitesimal**. *Se f é p vezes diferenciável no ponto a, então* $\lim\limits_{v\to 0} \dfrac{r_p(v)}{|v|^p} = 0$.

2) **Resto de Lagrange**. *Supondo $[a, a+v] \subset U$, f de classe C^p, $p+1$ vezes diferenciável no segmento aberto $(a, a+v)$, então existe $\theta \in (0,1)$ tal que*

$$r_p(v) = \frac{1}{(p+1)!}\, d^{p+1} f(a+\theta v)\cdot v^{p+1}.$$

3) **Resto Integral**. *Se f é de classe C^{p+1} e $[a, a+v] \subset U$ então*

$$r_p(v) = \frac{1}{p!} \int_0^1 (1-t)^p\, d^{p+1} f(a+tv)\cdot v^{p+1}\, dt.$$

Acima, para $v = (\alpha_1, \dots, \alpha_n)$, escrevemos

$$d^2 f(a)\cdot v^2 = \sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j}(a)\alpha_i\alpha_j\,,$$

$$d^3 f(a)\cdot v^3 = \sum_{i,j,k} \frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k}(a)\alpha_i\alpha_j\alpha_k\,,$$

e assim por diante.

Para cada inteiro $p > 0$, a forma $d^p f(a)\colon \mathbb{R}^n \to \mathbb{R}$ chama-se a *p-ésima diferencial* da função f no ponto a. O valor de $d^p f(a)$ num vetor $v \in \mathbb{R}^n$ é indicado pela notação $d^p f(a) \cdot v^p$, para sugerir que se trata de um polinômio homogêneo de grau p nas coordenadas de v. (Veja que $d^p f(a) \cdot (tv)^p = t^p \cdot d^p(a) \cdot v^p$, o que é coerente com a notação.)

As afirmações 2) e 3) acima se demonstram simplesmente introduzindo a função $\varphi\colon [0,1] \to \mathbb{R}$, definida por $\varphi(t) = f(a+tv)$, aplicando a φ o resultado correspondente para funções de uma variável real, (Vol. 1, pags. 285 e 330), onde

$$\varphi(t) = \varphi(0) + \varphi'(0) + \frac{1}{2}\varphi''(0) + \cdots + \frac{1}{p!}\varphi^{(p)}(0) + r_p\,,$$

notando que $\varphi(0) = f(a)$, $\varphi(1) = f(a+v)$ e que a i-ésima derivada de φ no ponto $t \in [0,1]$, em virtude da Regra da Cadeia, é $\varphi^{(i)}(t) = d^i f(a+tv)\cdot v^i$. Com efeito, dada $\varphi(t) = f(a+tv)$, temos sucessivamente (derivando cada igualdade anterior):

$$\varphi'(t) = \sum_i \frac{\partial f}{\partial x_i}(a+tv)\alpha_i = df(a+tv)\cdot v,$$

$$\varphi''(t) = \sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j}(a+tv)\alpha_i\alpha_j = d^2 f(a+tv)\cdot v^2, \text{ etc.}$$

Quanto a 1), embora ela não decorra da fórmula de Taylor infinitesimal na reta, sua demonstração se faz de modo análogo. (Veja Vol. 1, pág. 283.)

Com efeito, o resto é uma função $r_p\colon U_0 \to \mathbb{R}$, definida no aberto $U_0 = \{v \in \mathbb{R}^n; a+v \in U\}$ por

$$r_p(v) = f(a+v) - f(a) - df(a)\cdot v - \cdots - \frac{1}{p!}d^p f(a)\cdot v^p.$$

A função r_p é p vezes diferenciável no ponto 0 e se anula nesse ponto, juntamente com suas derivadas parciais de ordem $\leq p$. A fórmula de Taylor infinitesimal resulta então do seguinte

Lema. *Seja $r\colon U \to \mathbb{R}$ uma função p vezes diferenciável no ponto $0 \in U$. Se r, juntamente com todas as suas derivadas parciais de ordem $\leq p$, se anula no ponto 0 então $\lim_{v\to 0} \dfrac{r(v)}{|v|^p} = 0$.*

Demonstração: Usaremos indução em p. O resultado é válido para $p = 1$, em virtude da própria definição de derivada. Supondo-o válido para p, seja r uma função $p+1$ vezes diferenciável no ponto 0, com todas as derivadas parciais de ordem $\leq p+1$ nulas na origem. Então, para cada $i = 1, \ldots, n$, a função $\dfrac{\partial r}{\partial x_i} : U \to \mathbb{R}$ é p vezes diferenciável e goza de propriedade semelhante, com p em vez de $p+1$. Pela hipótese, de indução, temos $\lim\limits_{v \to 0} \dfrac{\partial r}{\partial x_i}(v)/|v|^p = 0$. Ora, o Teorema do Valor Médio nos diz que existe $\theta \in (0,1)$ tal que:

$$\frac{r(v)}{|v|^{p+1}} = \sum_{i=1}^{n} \frac{\dfrac{\partial r}{\partial x_i}(\theta v) \cdot \alpha_i}{|v|^{p+1}} = \sum_{i=1}^{n} \frac{\dfrac{\partial r}{\partial x_i}(\theta v)}{|v|^p} \cdot \frac{\alpha_i}{|v|},$$

onde $= (\alpha_1, \ldots, \alpha_n)$. Podemos sempre supor $\alpha_i/|v|$ menor do que ou igual a 1, em valor absoluto. (Basta tomar em $\mathbb{R}^n$ a norma da soma.) Segue-se então que $\lim\limits_{v \to 0} \dfrac{r(v)}{|v|^{p+1}} = 0$, o que conclui a demonstração do lema.

Observação: É interessante observar que, como no caso $n = 1$ (veja loc. cit.), a recíproca do lema acima é verdadeira. Por simplicidade, consideremos $p = 2$. Então $r \colon U \to \mathbb{R}$ é uma função 2 vezes diferenciável na origem, a qual pertence ao aberto $U \subset \mathbb{R}^n$. Supomos que $\lim\limits_{v \to 0} r(v)/|v|^2 = 0$ e queremos provar que r e todas as suas derivadas parciais de ordem ≤ 2 se anulam no ponto $v = 0$. Da hipótese decorre que

$$r(0) = \lim_{v \to 0} r(v) = \lim_{v \to 0} \frac{r(v)}{|v|^2} \cdot |v|^2 = 0$$

e que

$$\lim_{v \to 0} \frac{r(v)}{|v|} = \lim_{v \to 0} \frac{r(v)}{|v|^2} |v| = 0.$$

Como

$$\lim_{t \to 0} \frac{r(te_i)}{|t|} = \lim_{t \to 0} \frac{r(te_i)}{|te_i|} = 0,$$

concluimos que $\dfrac{\partial r}{\partial x_i}(0) = 0$ para $i = 1, \ldots, n$.

Mostraremos agora que todas as derivadas parciais de ordem 2 de r na origem são nulas. Para isso, usaremos a fórmula de Taylor infinitesimal, segundo a qual

$$r(v) = \frac{1}{2}\sum_{i,j} \frac{\partial^2 r}{\partial x_i \partial y_j}(0)\alpha_i\alpha_j + \rho(v)\cdot|v|^2,$$

onde $v = (\alpha_1, \dots, \alpha_n)$ e $\lim_{v\to 0} \rho(v) = 0$. Tomemos inicialmente $v = te_i$. Então

$$r(te_i) = \frac{1}{2}\frac{\partial^2 r}{\partial x_i^2}(0)\cdot t^2 + \rho(te_i)\cdot t^2.$$

Dividindo esta igualdade por t^2 e fazendo $t \to 0$, obtemos $\frac{\partial^2 r}{\partial x_i^2}(0) = 0$ para $i = 1, \dots, n$. Agora tomemos $v = te_i + te_j$ $(i \neq j)$. Obtemos (derivadas tomadas no ponto 0):

$$r(te_i + te_j) = \left(\frac{1}{2}\frac{\partial^2 r}{\partial x_i^2} + \frac{1}{2}\frac{\partial^2 r}{\partial x_j^2} + \frac{\partial^2 r}{\partial x_i \partial x_j}\right)t^2 + \rho(te_i + te_j)2t^2,$$

pois $|te_i + te_j|^2 = 2t^2$. Como $\partial^2 r/\partial x_i^2 = 0$ para todo i, temos:

$$r(te_i + te_j) = \frac{\partial^2 r}{\partial x_i \partial x_j}\cdot t^2 + \rho(te_i + te_j)2t^2.$$

Dividindo esta igualdade por t^2 e fazendo $t \to 0$, resulta $\frac{\partial^2 r}{\partial x_i \partial x_j}(0) = 0$.

Um importante (e imediato) corolário desta Observação é a *unicidade da fórmula de Taylor*. Para enunciá-la, convencionaremos a seguinte notação: dada uma função i-linear $\varphi\colon \mathbb{R}^n \times \cdots \times \mathbb{R}^n \to \mathbb{R}$ escreveremos $\varphi \cdot v^i$ para significar $\varphi(v, v, \dots, v)$, onde $v \in \mathbb{R}^n$. Então, se $f\colon U \to \mathbb{R}$ é p vezes diferenciável no ponto $a \in U$ e, para cada $i = 1, 2, \dots, p$, é dada uma função i-linear $\varphi_i\colon \mathbb{R}^n \times \cdots \times \mathbb{R}^n \to \mathbb{R}$, de tal modo que valha

$$f(a+v) = f(a) + \varphi_1\cdot v + \varphi_2\cdot v^2 + \cdots + \varphi_p\cdot v^p + r_p(v),$$

onde $\lim_{v\to 0} \frac{r_p(v)}{|v|^p} = 0$, então

$$\varphi_i\cdot v^i = \frac{1}{i!}\, d^i f(a)\cdot v^i$$

para todo $i = 1, \dots, p$.

No momento, nosso maior interesse reside na diferencial segunda $d^2f(a)$, chamada a *forma Hessiana* da função f no ponto a. Ela é uma forma quadrática, conforme a definição abaixo.

Uma *forma quadrática* $H\colon \mathbb{R}^n \to \mathbb{R}$ é uma função cujo valor num vetor $v = (\alpha_1, \dots, \alpha_n)$ é dado por $\sum\limits_{i,j=1}^{n} h_{ij}\alpha_i\alpha_j$, onde (h_{ij}) é uma matriz simétrica $n \times n$. Indica-se com a notação $H \cdot v^2$ o valor da forma quadrática H no vetor v. Desta maneira,

$$H \cdot v^2 = \sum_{i,j=1}^{n} h_{ij}\alpha_i\alpha_j\,.$$

Se $t \in \mathbb{R}$ então $H \cdot (tv)^2 = t^2 \cdot (H \cdot v^2)$.

A forma hessiana da função duas vezes diferenciável $f\colon U \to \mathbb{R}$ no ponto $x \in U$ será indicada com $H(x)$, ou $Hf(x)$ caso seja necessário ser mais explícito. Sabemos que $H(x) = d^2f(x)$, portanto

$$H(x) \cdot v^2 = \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(x)\alpha_i\alpha_j\,.$$

O Teorema de Schwarz garante que a matriz $\left(\dfrac{\partial^2 f}{\partial x_i \partial x_j}(x)\right)$, chamada *matriz Hessiana* de f no ponto x, é simétrica.

Usaremos agora a fórmula de Taylor infinitesimal para estudar os pontos críticos de uma função de classe C^2.

Dada uma função diferenciável $f\colon U \to \mathbb{R}$, um ponto $a \in U$ chama-se *ponto crítico* de f (ou ponto *singular*) quando $df(a) = 0$, isto é, $\dfrac{\partial f}{\partial x_1}(a) = \cdots = \dfrac{\partial f}{\partial x_n}(a) = 0$.

Diz-se que a função f tem um *máximo* (resp. *mínimo*) *local* no ponto $a \in U$ quando existe $\delta > 0$ tal que $|v| < \delta \Rightarrow f(a+v) \le f(a)$ (resp. $f(a) \le f(a+v)$). Por exemplo, as funções $f(x,y) = x^2 + y^2$ e $g(x,y) = x^2$ têm mínimo local na origem.

Se f tem um máximo local (ou mínimo local) no ponto a, então a é um ponto crítico de f. Com efeito, neste caso 0 é um ponto de máximo local (ou de mínimo local) para cada função $\varphi_i(t) = f(a + te_i)$, $i = 1, \dots, n$. Pelo Cálculo de uma variável, segue-se que $\dfrac{\partial f}{\partial x_i}(a) = \varphi_i'(0) = 0$ para $i = 1, \dots, n$. Logo a é um ponto crítico de f.

O ponto crítico a diz-se *não-degenerado* quando a matriz Hessiana nesse ponto é invertível, isto é, $\det\left(\dfrac{\partial^2 f}{\partial x_i \partial x_j}(a)\right) \neq 0$.

Teorema 4. *Seja $f\colon U \to \mathbb{R}$ duas vezes diferenciável. Todo ponto crítico não-degenerado $a \in U$ é um ponto crítico isolado.*

O Teorema 4 é uma conseqüência imediata do

Teorema 4a. *Seja $F = (f_1, \dots, f_n)\colon U \to \mathbb{R}^n$, onde cada $f_i\colon U \to \mathbb{R}$ $(1 \leq i \leq n)$ é diferenciável no ponto $a \in U \subset \mathbb{R}^n$. Se a matriz $n \times n$, $H = \left(\dfrac{\partial f_i}{\partial x_j}(a)\right)$, tem determinante $\neq 0$ então existe $\delta > 0$ tal que $0 < |x - a| < \delta \Rightarrow F(x) \neq F(a)$.*

Utilizaremos o

Lema. *Seja $H\colon \mathbb{R}^n \to \mathbb{R}^n$ uma transformação linear invertível. Existe $c > 0$ tal que $|H \cdot x| \geq c|x|$ para todo $x \in \mathbb{R}^n$.*

Demonstração: Seja $c = 1/|H^{-1}|$. Para todo $x \in \mathbb{R}^n$, temos $|x| = |H^{-1}(Hx)| \leq |H^{-1}|\,|Hx| = 1/c|Hx|$, donde $|Hx| \geq c|x|$.

Demonstração do Teorema 4a: A definição de diferenciabilidade nos dá, para cada $i = 1, \dots, n$, e $x \in U$:

$$f_i(x) = f_i(a) + \sum_{j=1}^{n} h_{ij} \cdot (x_j - a_j) + \rho_i(x)|x - a|, \text{ onde } \lim_{x \to a} \rho_i(x) = 0,$$

e $h_{ij} = (\partial f_i / \partial x_j)(a)$. Pondo $R(x) = (\rho_1(x), \dots, \rho_n(x))$, vem

$$F(x) = F(a) + H \cdot (x - a) + R(x) \cdot |x - a|, \text{ onde } \lim_{x \to a} R(x) = 0.$$

Tomemos c como no Lema. Existe $\delta > 0$ tal que $0 < |x - a| < \delta \Rightarrow |R(x)| < \dfrac{c}{2} \cdot$ Portanto, para todo $x \in U$ com $0 < |x - a| < \delta$, temos:

$$\begin{aligned}|F(x) - F(a)| &\geq |H \cdot (x - a)| - |R(x)|\,|x - a)| \geq \\ &\geq c|x - a| - \frac{c}{2}\,|x - a| = \frac{c}{2}\,|x - a| > 0,\end{aligned}$$

donde $F(x) \neq F(a)$, o que prova o Teorema 4a.

Para obter o Teorema 4, basta tomar $f_i = \dfrac{\partial f}{\partial x_i}$; então

$$\frac{\partial f_i}{\partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i} \cdot$$

Corolário 1. *O conjunto dos pontos críticos não-degenerados de uma função duas vezes diferenciável é enumerável.*

Com efeito, todo conjunto cujos pontos são isolados é enumerável (veja (11.4), Cap. I).

Corolário 2. *Se todos os pontos críticos de uma função $f\colon U \to \mathbb{R}$, duas vezes diferenciável, são não-degenerados então em cada compacto $K \subset U$ há apenas um número finito deles.*

Com efeito, se $x \in U$ é limite de uma seqüência de pontos críticos de uma função de classe C^1 definida em U então x é ponto crítico dessa função. Logo o conjunto C dos pontos críticos de f contidos em K é um subconjunto fechado de K e portanto compacto. Como cada um dos pontos de C é isolado, segue-se que C é finito.

Seja $H\colon \mathbb{R}^n \to \mathbb{R}$ uma forma quadrática, dada por $H \cdot v^2 = \Sigma\, h_{ij}\alpha_i\alpha_j$ para $v = (\alpha_1, \ldots, \alpha_n)$. Diremos que a forma H é *positiva* quando tivermos $H \cdot v^2 > 0$ para todo $v \neq 0$ em $\mathbb{R}^n$. Se for $H \cdot v^2 < 0$ para todo $v \neq 0$, diremos que H é uma forma quadrática *negativa.* Se uma forma quadrática for positiva ou negativa, diremos que ela é uma forma *definida.* Chamaremos H de forma quadrática *indefinida* quando existirem vetores $v, w \in \mathbb{R}^n$ tais que $H \cdot v^2 > 0$ e $H \cdot w^2 < 0$.

Exemplo 15. Dado um produto interno $\langle x, v\rangle$ em $\mathbb{R}^n$, a forma quadrática $H \cdot v^2 = \langle v, v\rangle$ é positiva enquanto $H \cdot v^2 = -\langle v, v\rangle$ é negativa. Por outro lado, para todo inteiro $i \in [1, n)$, a forma quadrática $H \cdot v^2 = \alpha_1^2 + \cdots + \alpha_i^2 - \alpha_{i+1}^2 - \cdots - \alpha_n^2$ é indefinida.

Se uma forma H é definida então sua matriz (h_{ij}) é necessariamente invertível. Isto resulta da seguinte observação: se indicarmos com $H_0\colon \mathbb{R}^n \to \mathbb{R}^n$ a transformação linear cuja matriz relativamente à base canônica é (h_{ij}), temos $H \cdot v^2 = \langle H_0 \cdot v, v\rangle$ (produto interno canônico de $\mathbb{R}^n$). Assim, H definida $\Rightarrow \langle H_0 \cdot v, v\rangle \neq 0$ para todo $v \neq 0 \Rightarrow H_0 \cdot v \neq 0$ para todo $v \neq 0 \Rightarrow H_0$ invertível.

Portanto, se a forma Hessiana de uma função de classe C^2, num ponto crítico a, é positiva ou negativa, o ponto crítico em questão é não-degenerado.

A relação entre ponto crítico e a forma Hessiana é estabelecida pelo

Teorema 5. *Sejam $f\colon U \to \mathbb{R}$ uma função de classe C^2, $a \in U$ um ponto crítico de f e H a forma quadrática Hessiana de f no ponto a. Então:*

1) *Se H é positiva, a é um ponto de mínimo local não-degenerado;*

2) *Se H é negativa, a é um ponto máximo local não-degenerado;*

3) *Se H é indefinida, a não é ponto de mínimo local nem de máximo local para f.*

Demonstração: Provemos 1). Seja $B(a;\delta) \subset U$. Para todo v, com $0 < |v| < \delta$, temos

$$(*) \qquad f(a+v) = f(a) + \frac{1}{2} H \cdot v^2 + r(v) = \\ = f(a) + \left[\frac{1}{2} H \cdot \left(\frac{v}{|v|}\right)^2 + \frac{r(v)}{|v|^2}\right] \cdot |v|^2$$

A função contínua $\frac{1}{2} H$ sendo positiva em todos os pontos da esfera unitária de $\mathbb{R}^n$, a qual é um conjunto compacto, existe $\varepsilon > 0$ tal que $\frac{1}{2} H \cdot u^2 \geq \varepsilon$ para todo $u \in \mathbb{R}^n$ de norma 1. Então $\frac{1}{2} H \cdot \left(\frac{v}{|v|}\right)^2 \geq \varepsilon$ para todo $v \neq 0$ em $\mathbb{R}^n$. Ora, pela Fórmula de Taylor infinitesimal, podemos tomar $\delta > 0$ tão pequeno que $0 < |v| < \delta \Rightarrow |r(v)|/|v|^2 < \varepsilon$. Isto significa que, para todo $v \neq 0$ com norma inferior a δ, a expressão dentro dos colchetes em (*) é positiva. Portanto $0 < |v| < \delta \Rightarrow f(a+v) > f(a)$, e a é um ponto de mínimo local para f. A afirmação 2) se prova de modo semelhante.

Quanto à terceira afirmação do teorema, se H é indefinida, existem vetores $v, w \in \mathbb{R}^n$ tais que $H \cdot v^2 > 0$ e $H \cdot w^2 < 0$. Então, para todo $t \neq 0$ real, temos $H \cdot (tv)^2 = t^2 \cdot H \cdot v^2 > 0$ e $H \cdot (tw)^2 < 0$. A igualdade (*) acima implica então que $f(a+tv) > f(a)$ e $f(a+tw) < f(a)$ para todo $t \neq 0$ suficientemente pequeno, logo a não é ponto de mínimo local nem de máximo local para f.

Exemplo 16. Seja $f\colon \mathbb{R}^{m+n} = \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$ definida por $f(x,y) = \langle x,x\rangle - \langle y,y\rangle$, onde $x \in \mathbb{R}^m$ e $y \in \mathbb{R}^n$. Então $\partial f/\partial x_i = 2x_i$ e $\partial f/\partial y_j = -2y_j$, logo $\operatorname{grad} f(x,y) = 2(x,-y)$. O único ponto crítico de f é, portanto, a origem. Em todo ponto $(x,y) \in \mathbb{R}^{m+n}$, a matriz Hessiana de f é uma matriz diagonal, cujos m primeiros elementos são iguais a 2 e os n últimos iguais a -2. A forma quadrática Hessiana é, pois, constante; ela é positiva se $n=0$, negativa se $m=0$ e indefinida quando $m \cdot n \neq 0$. Assim, a origem é um ponto de mínimo se $n=0$, e de máximo quando $m=0$. Para $m \cdot n \neq 0$, f não admite mínimo nem máximo local na

origem, que se chama então um *ponto de sela*, devido à forma do gráfico da função $f(x,y) = x^2 - y^2$.

Exemplo 17. Como vimos na parte final da demonstração do teorema acima, se a forma Hessiana H da função f no ponto crítico a é tal que $H \cdot v^2 > 0$ para algum $v \in \mathbb{R}^n$, então $f(a+tv) > f(a)$ para todo $t \neq 0$ suficientemente pequeno. Daí resulta que se a função f tem, num ponto a, um máximo local então sua forma quadrática Hessiana é, nesse ponto, não-positiva, isto é, $H \cdot v^2 \leq 0$ para todo $v \in \mathbb{R}^n$. Analogamente, se f tem um mínimo local no ponto a, sua forma Hessiana nesse ponto é ≥ 0. Mas as recíprocas destas afirmações são falsas: quando a forma Hessiana num ponto é apenas ≥ 0 (ou ≤ 0) não se pode afirmar que a função tem um mínimo (ou máximo) no ponto. Por exemplo, $f(x,y) = x^2$ e $g(x,y) = x^2 + y^3$ admitem a origem como ponto crítico, no qual estas funções têm a mesma forma Hessiana, $H \cdot v^2 = 2\alpha^2, \quad v = (\alpha, \beta)$. A origem é um ponto de mínimo para f mas não é um mínimo local para g.

O problema prático de determinar se uma forma quadrática H é ou não positiva (ou negativa) pode ser resolvido pelo método de Lagrange de completar quadrados. Este método se baseia na observação de que $a^2 + 2ab = (a+b)^2 - b^2$ e consiste em efetuar sucessivas mudanças de coordenadas que visam eliminar, na expressão de H, as parcelas como xy, xz, yz, etc., deixando apenas as do tipo x^2, y^2, z^2 etc.

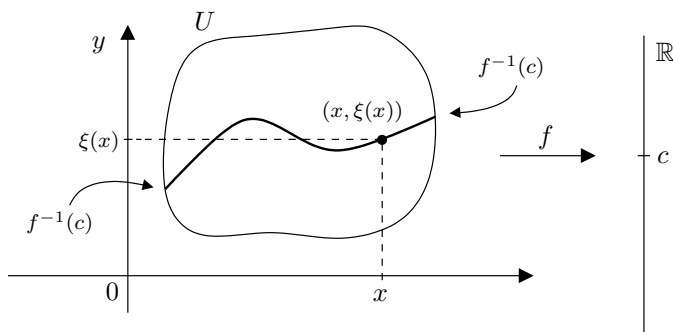
Exemplo 18. Seja em $\mathbb{R}^3$ a forma $H(x,y,z) = 3x^2 + 2y^2 - 5z^2 + 4xy + 2xz - yz$. A soma das parcelas que contêm x se escreve como $3x^2 + 4xy + 2xz = 3[x^2 + 2x\left(\frac{2}{3}y + \frac{1}{3}z\right)] = 3[\left(x + \frac{2}{3}y + \frac{1}{3}z\right)^2 - \left(\frac{2}{3}y + \frac{1}{3}z\right)^2] = 3u^2 - \frac{1}{3}(2y+z)^2$, onde $u = x + \frac{2}{3}y + \frac{1}{3}z$. Portanto $H(x,y,z) = 3u^2 - \frac{1}{3}(2y+z)^2 + 2y^2 - 5z^2 - yz = 3u^2 + \frac{2}{3}y^2 - \frac{16}{3}z^2 - \frac{7}{3}yz$. A soma das parcelas que contêm y nesta última expressão é $\frac{2}{3}y^2 - \frac{7}{3}yz = \frac{2}{3}\left(y^2 - 2y\frac{7}{4}z\right) = \frac{2}{3}[\left(y - \frac{7}{4}z\right)^2 - \frac{49}{16}z^2] = \frac{2}{3}v^2 - \frac{49}{24}z^2$, onde $v = y - \frac{7}{4}z$. Portanto

$$H(x,y,z) = 3u^2 + \frac{2}{3}v^2 - \frac{49}{24}z^2 - \frac{16}{3}z^2 = 3u^2 + \frac{2}{3}v^2 - \frac{177}{24}z^2.$$

Isto mostra que H é indefinida, pois assume valores positivos quando $z = 0$ e $u^2 + v^2 \neq 0$ e valores negativos, quando $u = v = 0$ e $z \neq 0$.

Observação: Se não houver, na expressão de H, termos em x^2, y^2, etc., a mudança de variáveis $x = s+t, \quad y = s-t$ faz $xy = s^2 - t^2$. Assim, por exemplo, $H(x,y,z) = xy + yz + xz = s^2 - t^2 + 2sz = s^2 + 2sz + z^2 - z^2 - t^2 = (s+z)^2 - z^2 - t^2 = u^2 - z^2 - t^2$, logo H é indefinida.

9 O teorema da função implícita

Por simplicidade, consideramos inicialmente funções de duas variáveis. Dada $f\colon U \to \mathbb{R}$, definida no aberto $U \subset \mathbb{R}^2$, e fixado $c \in \mathbb{R}$, dizemos que a equação $f(x,y) = c$ define y *implicitamente* como função de x quando existe uma função $\xi\colon I \to \mathbb{R}$, definida num intervalo $I \subset \mathbb{R}$, tal que $f(x,y) = c \Leftrightarrow y = \xi(x)$. Isto quer dizer que $f^{-1}(c)$ é o gráfico da função ξ.

Exemplo 19. É mais comum acontecer que uma equação do tipo $f(x,y) = c$ (quando define alguma coisa) defina y como função de x, ou x como função de y, *apenas localmente*. Por exemplo seja $f\colon \mathbb{R}^2 \to \mathbb{R}$ dada por $f(x,y) = x^2 + y^2$, e tomemos $c = 1$. A equação $x^2 + y^2 = 1$ não define y como função de x (nem x como função de y). [Por exemplo, para cada $x \in (-1,+1)$ existem 2 valores de y tais que $x^2+y^2 = 1$.] Mas, se tomarmos $U_1 = \{(x,y) \in \mathbb{R}^2; y > 0\}$, $U_2 = \{(x,y) \in \mathbb{R}^2; y < 0\}$, $U_3 = \{(x,y) \in \mathbb{R}^2; x > 0\}$ e $U_4 = \{(x,y) \in \mathbb{R}^2; x < 0\}$ então a equação $x^2 + y^2 = 1$ equivale a $y = \sqrt{1-x^2}$ quando $(x,y) \in U_1$, equivale a $y = -\sqrt{1-x^2}$ para $(x,y) \in U_2$, a $x = \sqrt{1-y^2}$ se $(x,y) \in U_3$ e a $x = -\sqrt{1-y^2}$ para $(x,y) \in U_4$. Como o círculo $S^1 = \{(x,y) \in \mathbb{R}^2; x^2 + y^2 = 1\}$, conjunto de todas as soluções (x,y) da equação $x^2 + y^2 = 1$, está contido na reunião dos 4 abertos U_1, U_2, U_3 e U_4, dizemos que a equação $x^2 + y^2 = 1$ define *localmente* y como função de x, ou x como função de y. Isto quer dizer que cada solução (x_0, y_0) desta equação está contida em algum aberto U_i tal que $f^{-1}(1) \cap U_i$ é o gráfico de uma função $[x = \xi(y)$, ou $y = \xi(x)]$.

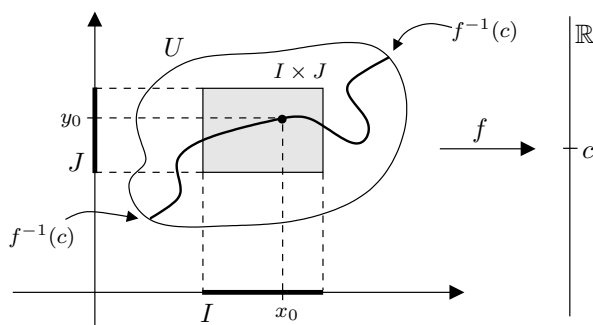
Notemos que é bem possível a uma equação do tipo $f(x,y) = c$ não definir função alguma: basta que c não pertença à imagem de f. Por exemplo, $x^2 + y^2 + 1 = 0$ não possui soluções reais (x,y), logo não define

y como função de x nem x como função de y. Mesmo que a equação $f(x,y) = c$ possua soluções, elas podem não definir funções, tal é o caso de $x^2 + y^2 = 0$: a única solução é $(0,0)$, que obviamente não é gráfico de uma função definida num intervalo não-degenerado. Outro exemplo elucidativo é o seguinte: a origem é solução da equação $x^2 - y^2 = 0$ mas, para nenhum aberto V contendo a origem, a interseção $f^{-1}(0) \cap V$ é o gráfico de uma função $y = \xi(x)$ ou $x = \zeta(y)$ pois tal interseção contém sempre 2 segmentos de reta de inclinação ± 1 que se cortam na origem.

Feitas essas considerações, enunciemos o

Teorema da Função Implícita. *Sejam* $f\colon U \to \mathbb{R}$ *uma função de classe* C^k $(k \geq 1)$, *definida num aberto* $U \subset \mathbb{R}^2$, *e* $(x_0, y_0) \in U$ *tal que* $f(x_0,y_0) = c$, $\dfrac{\partial f}{\partial y}(x_0,y_0) \neq 0$. *Então existe um retângulo aberto* $I \times J$, *de centro* (x_0,y_0), *tal que* $f^{-1}(c) \cap (I \times J)$ *é o gráfico de uma função* $\xi\colon I \to J$, *de classe* C^k. *Tem-se* $\xi'(x) = -\dfrac{\partial f/\partial x}{\partial f/\partial y}$ *estas derivadas sendo calculadas no ponto* $(x, \xi(x))$.

Como $(x_0,y_0) \in I \times J$, o intervalo aberto I contém x_0, enquanto J contém y_0.

A afirmação de que $f^{-1}(c) \cap (I \times J)$ é o gráfico de uma função $\xi\colon I \to J$ significa que, para cada $x \in I$, existe um único $y \in J$ com $f(x,y) = c$. Põe-se $y = \xi(x)$; a função $\xi\colon I \to J$ diz-se "definida implicitamente", no aberto $I \times J$, pela equação $f(x,y) = c$.

Demonstração: Para fixar as idéias, suponhamos $\dfrac{\partial f}{\partial y}(x_0,y_0) > 0$. Como $\dfrac{\partial f}{\partial y}$ é contínua, existem $\delta > 0$ e $\varepsilon > 0$ tais que, pondo $I = (x_0 - \delta, x_0 + \delta)$ e $J = (y_0 - \varepsilon, y_0 + \varepsilon)$, temos $I \times \overline{J} \subset U$ e $\dfrac{\partial f}{\partial y}(x,y) > 0$

para todo ponto $(x, y) \in I \times \overline{J}$. Então, para todo $x \in I$, a função $y \mapsto f(x, y)$ é estritamente crescente no intervalo $\overline{J}$. Em particular, como $f(x_0, y_0) = c$, temos $f(x_0, y_0 - \varepsilon) < c$ e $f(x_0, y_0 + \varepsilon) > c$. Pela continuidade de f, podemos supor δ tão pequeno que, para todo $x \in I$, tenhamos $f(x, y_0 - \varepsilon) < c$ e $f(x, y_0 + \varepsilon) > c$. Em virtude do Teorema do Valor Intermediário existe, para cada $x \in I$, um único $y = \xi(x) \in \overline{J}$ tal que $f(x, y) = c$. Tem-se obrigatoriamente $y \in J$, portanto $f^{-1}(c) \cap (I \times \overline{J}) = f^{-1}(c) \cap (I \times J)$ é o gráfico de uma função $\xi \colon I \to J$. Vamos mostrar que ξ é de classe C^k, ou seja, que existe $\xi'(x)$ para todo $x \in I$ e que $\xi' \colon I \to \mathbb{R}$ é de classe C^{k-1}.

Ora, pondo $k = \xi(x + h) - \xi(x)$, temos $\xi(x + h) = \xi(x) + k$, logo $f(x+h, \xi(x)+k) = f(x, \xi(x)) = c$. Pelo Teorema do Valor Médio, existe θ, com $0 < \theta < 1$, tal que

$$0 = f(x + h, \xi(x) + k) - f(x, \xi(x)) =$$
$$= \frac{\partial f}{\partial x}(x + \theta h, \xi(x) + \theta k) \cdot h + \frac{\partial f}{\partial y}(x + \theta h, \xi(x) + \theta k) \cdot k$$

Daí

$$\frac{\xi(x + h) - \xi(x)}{h} = \frac{k}{h} = -\frac{\dfrac{\partial f}{\partial x}(x + \theta h, \xi(x) + \theta k)}{\dfrac{\partial f}{\partial y}(x + \theta h, \xi(x) + \theta k)}.$$

Segundo o lema que provaremos logo a seguir, ξ é contínua. Isto significa que $\lim\limits_{h \to 0} k = 0$. A continuidade das derivadas parciais de f nos dá portanto

$$\xi'(x) = \lim_{h \to 0} \frac{\xi(x + h) - \xi(x)}{h} = -\frac{\dfrac{\partial f}{\partial x}(x, \xi(x))}{\dfrac{\partial f}{\partial y}(x, \xi(x))}.$$

Se $f \in C^1$, sendo $\partial f/\partial x$, $\partial f/\partial y$ e ξ contínuas, esta fórmula mostra que ξ' é contínua, logo $\xi \in C^1$. Se $f \in C^2$ então $\partial f/\partial x$, $\partial f/\partial y$ e (como acabamos de mostrar) ξ são de classe C^1. A fórmula que dá ξ' mostra então que ξ' é também de classe C^1, isto é, $\xi \in C^2$. E assim por diante: se $f \in C^k$ então $\xi \in C^k$.

Vejamos agora o lema usado na demonstração.

Lema. *Sejam $X \subset \mathbb{R}^m$, $K \subset \mathbb{R}^k$ compacto, $f \colon X \times K \to \mathbb{R}^p$ contínua e $c \in \mathbb{R}^p$. Se $f^{-1}(c)$ é o gráfico de uma aplicação $\xi \colon X \to K$,[isto é, para*

cada $x \in X$ *existe um único* $y = \xi(x) \in K$ *com* $f(x, \xi(x)) = c]$ *então* ξ *é contínua.*

Demonstração: Dado $x_0 \in X$, seja $y_0 = \xi(x_0)$. Tomamos uma seqüência de pontos $x_n \in X$, com $\lim x_n = x_0$, e queremos provar que $\lim \xi(x_n) = y_0$. Como a seqüência $(\xi(x_n))$ é limitada (pois $\xi(x_n) \in K$ para todo n), basta provar que toda subseqüência $\xi(x'_n)$, convergente em $\mathbb{R}^k$, tem limite y_0. Ora, se for $\lim \xi(x'_n) = y$, deve ser $y \in K$ pois K é fechado. Como $f(x'_n, \xi(x'_n)) = c$ para todo n, temos $f(x_0, y) = \lim f(x'_n, \xi(x'_n)) = c$. Pela unicidade de y_0, isto obriga $y = y_0$ e conclui a demonstração.

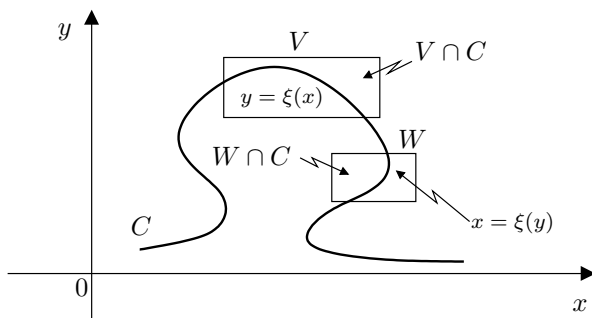
Contra-exemplo ao Lema, supondo apenas K limitado mas não compacto: seja $f : \mathbb{R} \times [0, 1) \to \mathbb{R}$ definida por $f(x, y) = (x^2 + y^2)(ye^{|x|} - 1)$. Então $f^{-1}(0)$ é o gráfico da função $\xi : \mathbb{R} \to [0, 1)$, onde $\xi(x) = e^{-|x|}$ se $x \neq 0$ e $\xi(0) = 0$. Embora f seja contínua, ξ é descontínua.

Para interpretar geometricamente o Teorema da Função Implícita, vamos introduzir algumas definições.

Seja $f : U \to \mathbb{R}$ uma função diferenciável no aberto $U \subset \mathbb{R}^n$. Diremos que o número real c é um *valor regular* de f quando não existirem pontos críticos de f no nível c, ou seja, $f(x) = c \Rightarrow \operatorname{grad} f(x) \neq 0$. (Note que esta definição está formulada de tal modo que $f^{-1}(c) = \varnothing$ implica automaticamente que c é um valor regular de f.) Quando c é um valor regular de f, diz-se também que o *nível* c *é regular*. Quando existem pontos críticos $x \in U$ tais que $f(x) = c$, dizemos que c *é um nivel crítico de* f.

Exemplo 20. Qualquer número real diferente de 3 é valor regular da função $f(x, y) = x^2 - y^2 + 3$ pois $\operatorname{grad} f(x, y) = (2x, -2y)$ só se anula no ponto $(x, y) = (0, 0)$, no qual o valor de f é 3. Por outro lado $f(0, 0) = 3$ não é valor regular de f pois $\operatorname{grad} f(0, 0) = (0, 0)$.

Um conjunto $C \subset \mathbb{R}^2$ chama-se uma *curva de classe* C^k $(k \geq 0)$ quando C é localmente o gráfico de uma função de classe C^k. Isto quer dizer que cada ponto $p \in C$ está contido num aberto V tal que $V \cap C$ é o gráfico de uma função de classe C^k.

Exemplo 21. Vimos no Exemplo 19 que o círculo S^1 é uma curva de classe C^∞. Seja agora $C = \{(x, y) \in \mathbb{R}^2; x^2 - y^2 = 1\}$ (hipérbole). Afirmamos que C é uma curva (desconexa) de classe C^∞. Com efeito, sejam $V_1 = \{(x, y) \in \mathbb{R}^2; x > 0\}$ e $V_2 = \{(x, y) \in \mathbb{R}^2; x < 0\}$. Todo ponto

$p \in C$ pertence a V_1 ou a V_2. Além disso, $V_1 \cap C$ é o gráfico da função $x = \sqrt{1+y^2}$ enquanto $V_2 \cap C$ é o gráfico de $x = -\sqrt{1+y^2}$. Ambas são funções de classe C^∞ na reta, logo C é uma curva de classe C^∞.

O enunciado geométrico do Teorema da Função Implícita é o seguinte:

Seja $f\colon U \to \mathbb{R}$ uma função de classe C^k $(k \geq 1)$ no aberto $U \subset \mathbb{R}^2$. Para todo valor regular c da função f, o conjunto $f^{-1}(c)$ (se não for vazio) é uma curva de classe C^k, chamada a curva de nível *c da função f.*

Quando c não é um valor regular de f, a imagem inversa $f^{-1}(c)$ pode ou não ser uma curva. Por exemplo, se $f\colon \mathbb{R}^2 \to \mathbb{R}$ é dada por $f(x,y) = x^2 - y^2 + 3$ então $f^{-1}(3)$, reunião de duas retas que se cortam na origem, não é uma curva. Mas se $g(x,y) = x^2$ então 0 não é um valor regular de g mas $g^{-1}(0) =$ eixo dos y, é uma curva.

Trataremos agora de funções implícitas com um número qualquer de variáveis.

No teorema abaixo, representaremos os pontos de $\mathbb{R}^{n+1}$ por pares (x,y), onde $x \in \mathbb{R}^n$ e $y \in \mathbb{R}$.

Teorema da Função Implícita. *Seja $f\colon U \to \mathbb{R}$ uma função de classe C^k $(k \geq 1)$, definida num aberto $U \subset \mathbb{R}^{n+1}$. Se um ponto $p = (x_0, y_0) \in U$ é tal que $f(p) = c$ e $\dfrac{\partial f}{\partial y}(p) \neq 0$, então existem uma bola $B = B(x_0, \delta) \subset \mathbb{R}^n$ e um intervalo $J = (y_0 - \varepsilon, y_0 + \varepsilon)$ tais que $f^{-1}(c) \cap (B \times J)$ é o gráfico de uma função $\xi\colon B \to J$, de classe C^k.*

Para todo $x \in B$, tem-se

$$\frac{\partial \xi}{\partial x_i}(x) = -\frac{\partial f}{\partial x_i}(x, \xi(x)) \Big/ \frac{\partial f}{\partial y}(x, \xi(x)), \quad (i = 1, \ldots, n).$$

A função $y = \xi(x)$ diz-se "definida implicitamente pela equação $f(x, y) = c$". A afirmação de que $f^{-1}(c) \cap (B \times J)$ é o gráfico de uma função significa que, para cada $x \in B$ existe um único $y = \xi(x) \in J$ tal que $f(x, y) = c$. Evidentemente, $\xi(x_0) = y_0$.

Demonstração: Não há diferença essencial entre o caso geral e o já demonstrado, em que $n = 1$. Vamos, entretanto, repetir a prova, por cortesia. Para fixar idéais, seja $\frac{\partial f}{\partial y}(x_0, y_0) > 0$. Como $\partial f / \partial y$ é contínua, existem $\delta > 0$, $\varepsilon > 0$ tais que, pondo $B = B(x_0; \delta)$ e $J = (y_0 - \varepsilon, y_0 + \varepsilon)$, temos $B \times \overline{J} \subset U$ e $\frac{\partial f}{\partial y}(x, y) > 0$ para todo ponto $(x, y) \in B \times \overline{J}$. Então, para todo $x \in B$, a função $y \mapsto f(x, y)$ é estritamente crescente no intervalo $\overline{J} = [y_0 - \varepsilon, y_0 + \varepsilon]$. Como $f(x_0, y_0) = c$, segue-se que $f(x_0, y_0 - \varepsilon) < c$ e $f(x_0, y_0 + \varepsilon) > c$. Pela continuidade de f, podemos supor δ tão pequeno que, para todo $x \in B$, tenhamos $f(x, y_0 - \varepsilon) < c$ e $f(x, y_0 + \varepsilon) > c$. Em virtude do Teorema do Valor Intermediário existe, para cada $x \in B$, um único $y = \xi(x) \in \overline{J}$ tal que $f(x, y) = c$. Tem-se obrigatoriamente $y \in J$, logo $f^{-1}(c) \cap (B \times \overline{J}) = f^{-1}(c) \cap (B \times J)$ é o gráfico de uma função $\xi \colon B \to J$ a qual, pelo lema anterior, é contínua. Mostraremos agora que, em todo ponto $x \in B$, existem as derivadas parciais de ξ. Com efeito, pondo $k = k(t) = \xi(x + te_i) - \xi(x)$, temos $\xi(x + te_i) = \xi(x) + k$, logo $f(x + te_i, \xi(x) + k) = f(x, \xi(x)) = c$ para todo $t \in (-\delta, \delta)$.

Pelo Teorema do Valor Médio, para todo $t \in (-\delta, \delta)$ existe $\theta = \theta(t) \in (0, 1)$ tal que

$$0 = f(x + te_i, \xi(x) + k) - f(x, \xi(x)) =$$
$$= \frac{\partial f}{\partial x_i}(x + \theta te_i, \xi(x) + \theta k) \cdot t + \frac{\partial f}{\partial y}(x + \theta te_i, \xi(x) + \theta k) \cdot k.$$

Logo

$$\frac{\xi(x + te_i) - \xi(x)}{t} = \frac{k}{t} = -\frac{\frac{\partial f}{\partial x_i}(x + \theta te_i, \xi(x) + \theta k)}{\frac{\partial f}{\partial y}(x + \theta te_i, \xi(x) + \theta k)} \cdot$$

Pela continuidade de ξ, $\lim_{t\to 0} k(t) = 0$. A continuidade das derivadas parciais de f nos dá então

$$\frac{\partial \xi}{\partial x_i}(x) = \lim_{t\to 0} \frac{\xi(x+te_i) - \xi(x)}{t} = -\frac{\dfrac{\partial f}{\partial x_i}(x,\xi(x))}{\dfrac{\partial f}{\partial y}(x,\xi(x))}\,.$$

Sendo $f \in C^1$, resulta desta fórmula que as derivadas parciais de ξ são contínuas, logo $\xi \in C^1$. Se for $f \in C^2$, então suas derivadas parciais são de classe C^1. Como já temos $\xi \in C^1$, resulta ainda da fórmula acima que as derivadas parciais de ξ são de classe C^1, logo $\xi \in C^2$. E assim por diante: $f \in C^k$ implica $\xi \in C^k$.

Corolário. *Seja $f\colon U \to \mathbb{R}$ de classe $C^k\,(k \geq 1)$ no aberto $U \subset \mathbb{R}^{n+1}$. Se $\xi\colon V \to \mathbb{R}$ é contínua no aberto $V \subset \mathbb{R}^n$ com $(x,\xi(x)) \in U$, $\dfrac{\partial f}{\partial y}(x,\xi(x)) \neq 0$ e $f(x,\xi(x)) = c$ para todo $x \in V$, então ξ é de classe C^k.*

Evidentemente, não há nada especial a respeito da última variável. Dada $f\colon U \to \mathbb{R}$, de classe C^k no aberto $U \subset \mathbb{R}^{n+1}$, se num ponto $p \in U$ tivermos $f(p) = c$ e $\dfrac{\partial f}{\partial x_i}(p) \neq 0$ para algum inteiro $i \in [1, n+1]$ então existe um aberto $V \subset \mathbb{R}^{n+1}$, contendo p, tal que $f^{-1}(c) \cap V$ é o gráfico de uma função de n variáveis, de classe C^k. Mais precisamente, as condições $x \in V$, $\;f(x) = c$ definem, de modo único, a i-ésima coordenada de x em função das n restantes: $x_i = \xi(x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_{n+1})$, sendo ξ uma função de classe C^k.

Um conjunto $M \subset \mathbb{R}^{n+1}$ chama-se uma *hiperfície de classe C^k* quando é localmente o gráfico de uma função de n variáveis de classe C^k. Isto significa que cada ponto $p \in M$ pertence a um aberto $V \subset \mathbb{R}^{n+1}$ tal que $V \cap M$ é o gráfico de uma função de classe C^k definida num aberto do espaço $\mathbb{R}^n$. Quando $n = 1$ diz-se "curva"e, se $n = 2$, diz-se "superfície"em vez de "hiperfície".

A definição acima inclui o caso $k = 0$. As hiperfícies de classe C^0 têm interesse apenas topológico, não possuindo propriedades diferenciais. Podemos considerar também as *hiperfícies diferenciáveis* (caso intermediário entre C^0 e C^1), que são localmente gráficos de funções diferenciáveis. Evidentemente, toda hiperfície de classe C^k $(k \geq 1)$ é diferenciável.

Exemplo 22. Seja $S^n = \{x \in \mathbb{R}^{n+1}; \langle x, x\rangle = 1\}$ a esfera unitária n-dimensional. Indiquemos com $U \subset \mathbb{R}^n$ a bola aberta de raio 1, com centro na origem. Para cada $i = 1, \ldots, n+1$ ponhamos $V_i = \{x \in \mathbb{R}^{n+1}; x_i > 0\}$ e $W_i = \{x \in \mathbb{R}^{n+1}; x_i < 0\}$.

Escrevendo $x^* = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_{n+1})$, temos:

$$x \in S^n \cap V_i \Leftrightarrow |x^*| < 1 \quad \text{e} \quad x_i = \sqrt{1 - \langle x^*, x^*\rangle};$$
$$x \in S^n \cap W_i \Leftrightarrow |x^*| < 1 \quad \text{e} \quad x_i = -\sqrt{1 - \langle x^*, x^*\rangle}.$$

Logo, se considerarmos a função $\xi\colon U \to \mathbb{R}$, de classe C^∞, definida por $\xi(u) = \sqrt{1 - \langle u, u\rangle}$, vemos que, para cada $i = 1, \ldots, n+1$, $S^n \cap V_i$ é o gráfico da função $x_i = \xi(x^*)$ enquanto que $S^n \cap W_i$ é o gráfico de $x_i = -\xi(x^*)$. Como todo ponto $p \in S^n$ está contido em algum V_i ou em algum W_i, concluimos que S^n é uma hiperfície de classe C^∞ em $\mathbb{R}^{n+1}$.

Seja $M \subset \mathbb{R}^{n+1}$. Dado $p \in M$, usaremos a notação T_pM para indicar o conjunto dos vetores velocidade $\lambda'(0)$, dos caminhos $\lambda\colon (-\varepsilon, \varepsilon) \to M \subset \mathbb{R}^{n+1}$, contidos em M, diferenciáveis no ponto $t = 0$ e tais que $\lambda(0) = p$. Quando M é uma hiperfície diferenciável, o conjunto T_pM chama-se o *espaço vetorial tangente a M no ponto p*. Esta denominação tem sua justificativa no

Teorema 6. *Se a hiperfície $M \subset \mathbb{R}^{n+1}$ é diferenciável então, para cada $p \in M$, o conjunto T_pM é um subespaço vetorial de dimensão n do espaço euclidiano $\mathbb{R}^{n+1}$.*

Demonstração: Dado $p = (a_1, \ldots, a_{n+1})$ em M, existem abertos $V \subset \mathbb{R}^{n+1}$, $U \subset \mathbb{R}^n$, com $p \in V$, um inteiro $i \in [1, n+1]$ e uma função $\xi\colon U \to \mathbb{R}$, diferenciável, tais que $x \in V \cap M \Leftrightarrow x_i = \xi(x^*)$, onde $x^* = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_{n+1}) \in U$. (Assim nenhuma das variáveis de ξ tem índice i.) Escrevendo também $p^* = (a_1, \ldots, a_{i-1}, a_{i+1} \ldots, a_{n+1})$, afirmamos que

$$(*) \qquad v = (\alpha_1, \ldots, \alpha_{n+1}) \in T_pM \Leftrightarrow \alpha_i = \sum_{j \neq i} \frac{\partial \xi}{\partial x_j}(p^*) \cdot \alpha_j\,.$$

De fato, em primeiro lugar, se $v \in T_pM$ então $v = \lambda'(0)$ onde, restringindo ε se necessário, podemos supor que $\lambda\colon (-\varepsilon, \varepsilon) \to M \cap V$, $\lambda(0) = p$. Então, para todo $t \in (-\varepsilon, \varepsilon)$ temos

$$\lambda_i(t) = \xi(\lambda_1(t), \ldots, \lambda_{i-1}(t), \lambda_{i+1}(t), \ldots, \lambda_{n+1}(t)).$$

Pela Regra da Cadeia,

$$\lambda_i'(0) = \sum_{j \neq i} \frac{\partial \xi}{\partial x_j}(p^*) \cdot \lambda_j'(0), \text{ isto é, } \alpha_i = \sum_{j \neq i} \frac{\partial \xi}{\partial x_j}(p^*) \cdot \alpha_j\,,$$

como foi afirmado.

Reciprocamente, se o vetor $v = (\alpha_1, \ldots, \alpha_{n+1})$ é tal que $\alpha_i = \sum_{j \neq i} \frac{\partial \xi}{\partial x_j}(p^*)\alpha_j$ então definimos um caminho $\lambda\colon (-\varepsilon, \varepsilon) \to M \cap V$ tomando $\varepsilon > 0$ tal que $|t| < \varepsilon \Rightarrow p^* + tv^* \in U$, pondo $\lambda_j(t) = a_j + t\alpha_j$ para $j \neq i$ e

$$\begin{aligned}\lambda_i(t) = \xi(a_1 + t\alpha_1, \ldots, a_{i-1} + t\alpha_{i-1}, a_{i+1} + t\alpha_{i+1}, \ldots, a_{n+1} + \\ + t\alpha_{n+1}) = \xi(p^* + tv^*).\end{aligned}$$

Pela Regra da Cadeia, temos $\lambda'(0) = v$, logo $v \in T_pM$. A caracterização (*) acima obtida mostra que T_pM é um subespaço vetorial de dimensão n de $\mathbb{R}^{n+1}$, gerado pelos n vetores linearmente independentes

$$e_1 + c_1e_i, \ldots, e_{i-1} + c_{i-1}e_i, e_{i+1} + c_{i+1}e_i, \ldots, e_{n+1} + c_{n+1}e_i\,,$$

onde $c_j = (\partial\xi/\partial x_j)(p^*)$. Outra maneira de interpretar a afirmação (*) é dizer que ela caracteriza T_pM como o núcleo do funcional linear não nulo $\varphi\colon \mathbb{R}^{n+1} \to \mathbb{R}$, dado por $\varphi(v) = \alpha_i - \sum_{j \neq i} c_j\alpha_j$, onde $v = (\alpha_1, \ldots, \alpha_{n+1})$ e c_j é a j-ésima derivada parcial de ξ no ponto p^*. Ou ainda: T_pM é o gráfico do funcional linear $d\xi(p^*)$.

Exemplo 23. Para hiperfícies $M \subset \mathbb{R}^{n+1}$ de classe C^0, T_pM pode não ser um espaço vetorial de dimensão n; se $X = \{(x, y, z) \in \mathbb{R}^3; z = \sqrt{x^2 + y^2}\}$ (cone), $Y = \{(x, y, z) \in \mathbb{R}^3; z = |x|\}$ e $p = (0, 0, 0)$ então T_pX consiste do vetor 0 apenas, enquanto T_pY é o subespaço vetorial de dimensão 1 em $\mathbb{R}^3$, formado pelos vetores $(0, \beta, 0)$, $\beta \in \mathbb{R}$.

Exemplo 24. Sabemos que a esfera $S^n \subset \mathbb{R}^{n+1}$ é uma hiperfície de classe C^∞. Para cada ponto $p \in S^n$, o espaço vetorial tangente $T_p(S^n)$ é o conjunto $[p]^\perp$ dos vetores $v \in \mathbb{R}^{n+1}$ tais que $\langle p, v\rangle = 0$. Com efeito, se $\lambda\colon (-\varepsilon, \varepsilon) \to S^n$ é um caminho diferenciável no ponto $t = 0$, com $\lambda(0) = p$ então, como $|\lambda(t)| = 1$ para todo t, temos $\langle\lambda'(0), \lambda(0)\rangle = 0$, isto é, $\lambda'(0) \in [p]^\perp$. Isto mostra que $T_pM \subset [p]^\perp$. Como $[p]^\perp$ é um subespaço n-dimensional de $\mathbb{R}^{n+1}$, resulta que esta inclusão é, na realidade, uma igualdade.

O teorema abaixo permite obter um grande número de exemplos de hiperfícies. Ele diz que se não há pontos críticos de f no nível c então a "superfície de nível c" da função f é, de fato, uma hiperfície.

Teorema Global da Função Implícita. *A imagem inversa* $M = f^{-1}(c)$ *de um valor regular c de uma função* $f: U \to \mathbb{R}$, *de classe* C^k $(k \geq 1)$ *num aberto* $U \subset \mathbb{R}^{n+1}$, *é uma hiperfície de classe* C^k. *Em cada ponto* $p \in M$, *o espaço vetorial tangente* T_pM *é o núcleo da diferencial* $df(p)\colon \mathbb{R}^{n+1} \to \mathbb{R}$ *ou, equivalentemente, o conjunto dos vetores* $v \in \mathbb{R}^{n+1}$ *perpendiculares ao vetor* $\operatorname{grad} f(p)$.

Demonstração: O fato de que $M = f^{-1}(c)$ é uma hiperfície de classe C^k decorre imediatamente do Teorema da Função Implícita. Quanto a T_pM, se o caminho $\lambda\colon (-\varepsilon, \varepsilon) \to M$, com $\lambda(0) = p$, é diferenciável no ponto $t = 0$ então, sendo $f(\lambda(t)) = c$ para todo $t \in (-\varepsilon, \varepsilon)$, segue-se que $0 = (f \circ \lambda)'(0) = df(p) \cdot \lambda'(0) = \langle \operatorname{grad} f(p), \lambda'(0) \rangle$. Assim T_pM está contido no conjunto dos vetores de $\mathbb{R}^{n+1}$ que são perpendiculares a $\operatorname{grad} f(p)$. Como este conjunto é também subespaço vetorial de dimensão n, segue-se que ele coincide com T_pM.

Exemplo 25. Reexaminemos a esfera S^n à luz do teorema acima. Considerando a função $f\colon \mathbb{R}^{n+1} \to \mathbb{R}$ (de classe C^∞), definida por $f(x) = \langle x, x \rangle$, temos $\dfrac{\partial f}{\partial x_i}(x) = 2x_i$, logo $\operatorname{grad} f(x) = 2x$ para todo $x \in \mathbb{R}^{n+1}$. Assim, $\operatorname{grad} f(x) = 0 \Leftrightarrow x = 0 \Leftrightarrow f(x) = 0$. Em particular, 1 é um valor regular de f, donde $S^n = f^{-1}(1)$ é uma hiperfície em $\mathbb{R}^{n+1}$. Para cada $p \in S^n$, como $\operatorname{grad} f(p) = 2p$, temos $T_p(S^n) = \{v \in \mathbb{R}^{n+1}; \langle v, 2p \rangle = 2\langle v, p \rangle = 0\} = [p]^\perp$.

Exemplo 26. A função $\det\colon \mathbb{R}^{n^2} \to \mathbb{R}$, que associa a cada matriz $n \times n$, $X = (x_{ij})$, o seu determinante, é de classe C^∞. A expansão de $\det . X$ segundo os elementos da i-ésima linha nos dá $\det . X = \sum\limits_{j=1}^{n} (-1)^{i+j} x_{ij} X_{[ij]}$, onde $X_{[ij]}$ é o determinante da matriz $(n-1) \times (n-1)$ que se obtém omitindo a i-ésima linha e a j-ésima coluna de X. Logo $\dfrac{\partial \det}{\partial x_{ij}}(X) = (-1)^{i+j} X_{[ij]}$, para cada $X \in \mathbb{R}^{n^2}$. Em particular, no ponto $X = I$ (= matriz identidade $n \times n$), temos $\dfrac{\partial \det}{\partial x_{ij}}(I) = \delta_{ij}$ ("delta de Kronecker": $\delta_{ij} = 0$ se $i \neq j$ e $\delta_{ii} = 1$). Isto significa que o gradiente da função det no ponto I é a matriz identidade. Seja agora $U \subset \mathbb{R}^{n^2}$

o subconjunto aberto formado pelas matrizes $n \times n$ que têm determinante $\neq 0$ (matrizes invertíveis). Então a restrição $\det\colon U \to \mathbb{R}$ é uma função sem pontos críticos. [Com efeito, basta que, para um dos índices i ou j fixado, todos os números $X_{[ij]}$ sejam nulos para que ocorra $\det X = \sum_i (-1)^{i+j} x_{ij} X_{[ij]} = \sum_j (-1)^{i+j} x_{ij} X_{[ij]} = 0$, logo $X \notin U$.] Assim, todo número real é um valor regular para a função $\det\colon U \to \mathbb{R}$. Em particular, $M = (\det)^{-1}(1)$ = conjunto das matrizes $n \times n$ que têm determinante igual a 1 é uma hiperfície em $\mathbb{R}^{n^2}$. [M é um grupo relativamente à multiplicação de matrizes, conhecido como o "grupo unimodular"de $\mathbb{R}^n$.] O espaço vetorial tangente $T_I M$, no ponto I (matriz identidade) consiste nas matrizes $X = (x_{ij})$ que são perpendiculares ao gradiente de det nesse ponto, o qual sabemos que é $I = (\delta_{ij})$. Ora, temos $\langle X, I \rangle = \sum_{i,j} x_{ij} \delta_{ij} = \sum_i x_{ii}$ = traço de X. Assim, o espaço vetorial tangente a M no ponto I é o conjunto das matrizes de traço nulo.

Observação: Toda hiperfície $M \subset \mathbb{R}^{n+1}$, sendo localmente o gráfico de uma função $x_i = \xi(x, \ldots, x_{i-1}, x_{i+1}, \ldots, x_{n+1}) = \xi(x^*)$, de n variáveis, é também localmente a imagem inversa $f^{-1}(0)$ do valor regular 0 pela função $f(x) = x_i - \xi(x^*)$, definida no aberto V tal que $M \cap V$ é o gráfico de ξ. Não é verdade porém que toda hiperfície $M \subset \mathbb{R}^{n+1}$ seja globalmente do tipo $M = f^{-1}(c)$, imagem inversa de um valor regular, como no teorema acima. Se M é deste tipo então a aplicação contínua $\varphi = \operatorname{grad} f\colon M \to \mathbb{R}^{n+1}$ fornece o que se chama um "campo contínuo de vetores normais não nulos ao longo de M". Aqui, *normal* significa que, para cada $p \in M$, $\varphi(p) = \operatorname{grad} f(p)$ é perpendicular a todo vetor tangente $v \in T_p M$. As hiperfícies que admitem um campo contínuo de vetores normais não nulos $\varphi\colon M \to \mathbb{R}^{n+1}$ chamam-se hiperfícies *orientáveis*. Um exemplo de superfície não-orientável em $\mathbb{R}^3$ é a faixa de Möbius. (Veja §14 do Capítulo V.)

10 Multiplicador de Lagrange

Sejam $M \subset \mathbb{R}^{n+1}$ uma hiperfície de classe C^k $(k \geq 1)$, contida num aberto $U \subset \mathbb{R}^{n+1}$, e $f\colon U \to \mathbb{R}$ uma função real de classe C^k. Pretendemos determinar os máximos locais, mínimos locais e, mais geralmente, os pontos críticos da restrição $f|M$. Antes de mais nada precisamos definir o que entendemos por um ponto crítico de $f|M$.

Os pontos críticos de f em U são, como sabemos, os pontos $x \in U$ tais que $\operatorname{grad} f(x) = 0$, isto é, $\dfrac{\partial f}{\partial v}(x) = 0$ para todo $v \in \mathbb{R}^{n+1}$. Isto quer dizer que, para todo caminho diferenciável $\lambda\colon (-\varepsilon, \varepsilon) \to \mathbb{R}^{n+1}$ com $\lambda(0) = x$, temos $(f \circ \lambda)'(0) = 0$. Usando este fato como guia, definiremos um *ponto crítico* de $f|M$ como um ponto $p \in M$ tal que $(f \circ \lambda)'(0) = 0$ para todo caminho diferenciável $\lambda\colon (-\varepsilon, \varepsilon) \to M$, com $\lambda(0) = p$. Isto significa que $\dfrac{\partial f}{\partial v}(p) = 0$ para todo $v \in T_pM$, ou seja, $p \in M$ é um ponto crítico da restrição $f|M$ se, e somente se, $\langle \operatorname{grad} f(p), v\rangle = 0$ para todo $v \in T_pM$, ou ainda, se, e somente se, o vetor $\operatorname{grad} f(p)$ é normal à hiperfície M no ponto p.

Se $p \in M$ é um ponto de máximo local (ou de mínimo local) para a restrição $f|M$ então, para todo caminho diferenciável $\lambda\colon (-\varepsilon, \varepsilon) \to M$, com $\lambda(0) = p$, 0 é um ponto de máximo (ou de mínimo) local para a função real $f \circ \lambda\colon (-\varepsilon, \varepsilon) \to \mathbb{R}$, logo $(f \circ \lambda)'(0) = 0$, portanto p é um ponto crítico de $f|M$ de acordo com a definição acima.

Evidentemente, os pontos críticos de f em U, que por acaso pertençam a M, se existirem, serão pontos críticos de $f|M$. O interessante, porém, é que podem ocorrer pontos críticos de $f|M$ que não são pontos críticos de f, isto é, nos quais $\operatorname{grad} f$ não se anula. Por exemplo, sejam $f\colon \mathbb{R}^2 \to \mathbb{R}$, dada por $f(x, y) = y$, e $M = S^1 = \{(x, y) \in \mathbb{R}^2; x^2+y^2 = 1\}$. Evidentemente, f não possui pontos críticos, pois $\operatorname{grad} f = (0, 1)$ em todos os pontos $(x, y) \in \mathbb{R}^2$. Mas $p = (0, 1)$ e $q = (0, -1)$ são pontos críticos de $f|S^1$, inclusive porque são os pontos onde $f|S^1$ atinge, respectivamente, seu máximo e seu mínimo. De um modo geral, se a hiperfície $M \subset \mathbb{R}^{n+1}$ é compacta, então $f|M$ admite pelo menos dois pontos críticos, a saber, os pontos onde $f|M$ assume seus valores máximo e mínimo.

A resposta ao nosso problema é dada pelo

Teorema do Multiplicador de Lagrange. *Consideremos $f\colon U \to \mathbb{R}$, uma função de classe C^k $(k \geq 1)$ no aberto $U \subset \mathbb{R}^{n+1}$, e $M = \varphi^{-1}(c)$ uma hiperfície contida em U, imagem inversa do valor regular $c \in \mathbb{R}$ por uma função $\varphi\colon U \to \mathbb{R}$, de classe C^k. Um ponto $p \in M$ é ponto crítico de $f|M$ se, e somente se, existe um número real λ tal que* $\operatorname{grad} f(p) = \lambda \cdot \operatorname{grad} \varphi(p)$.

Demonstração: Para todo ponto $p \in M$, temos $\operatorname{grad} \varphi(p) \perp T_pM$ pois M é superfície de nível de φ. Além disso p é ponto crítico de $f|M$ se, e somente se, $\operatorname{grad} f(p) \perp T_pM$. Como $T_pM \subset \mathbb{R}^{n+1}$ é um subespaço

vetorial de dimensão n, segue-se que $p \in M$ é ponto crítico de $f|M$ se, e somente se, $\operatorname{grad} f(p)$ é um múltiplo de $\operatorname{grad} \varphi(p)$.

Observação: Quando a hiperfície M não é dada como imagem inversa $\varphi^{-1}(c)$ de um valor regular, os pontos críticos de $f|M$ são simplesmente os pontos $p \in M$ nos quais $\operatorname{grad} f(p)$ é normal a M.

A pesquisa dos pontos críticos de $f|M$ reduz-se, portanto, a resolver o sistema de $n+2$ equações

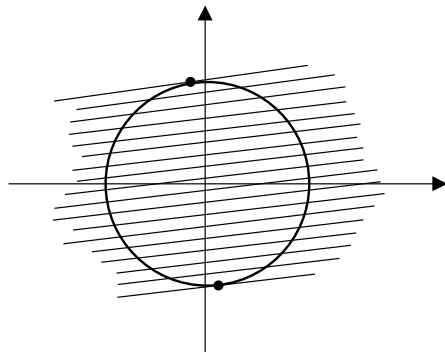
$$\begin{cases} \dfrac{\partial f}{\partial x_i}(p) = \lambda \cdot \dfrac{\partial \varphi}{\partial x_i}(p), \quad (i = 1, 2, \ldots, n+1); \\ \varphi(p) = c, \end{cases}$$

nas $n+2$ incógnitas λ, $x_1, \ldots, x_{n+1}$, onde $p = (x_1, \ldots, x_{n+1})$.

O número λ chama-se o *multiplicador de Lagrange*. Sua presença no sistema acima torna o número de equações igual ao número de incógnitas, o que muitas vezes ajuda a resolvê-lo.

A condição $\operatorname{grad} f(p) = \lambda \cdot \operatorname{grad} \varphi(p)$ significa que a hiperfície M é tangente à superfície de nível de f que passa pelo ponto crítico p da função $f|M$. No caso em que se podem esboçar essas superfícies, esta observação auxilia a localizar os pontos críticos.

Vejamos agora alguns problemas onde se aplica o método do multiplicador de Lagrange.

Exemplo 27. Seja $f\colon \mathbb{R}^2 \to \mathbb{R}$ definida por $f(x,y) = ax + by$, com $a^2 + b^2 \neq 0$. Quais são os pontos críticos da restrição de f ao círculo unitário S^1? Temos $\operatorname{grad} f = (a,b)$, $S^1 = \varphi^{-1}(1)$, $\varphi(x,y) = x^2 + y^2$ e $\operatorname{grad} \varphi = (2x, 2y)$. Os pontos críticos de $f|S^1$ são aqueles onde os vetores (a,b) e $(2x,2y)$ são colineares. Como, além disso, deve-se ter

$x^2+y^2=1$, isto nos dá $x=\dfrac{a}{\sqrt{a^2+b^2}}$, $y=\dfrac{b}{\sqrt{a^2+b^2}}$ ou $x=\dfrac{-a}{\sqrt{a^2+b^2}}$, $y=\dfrac{-b}{\sqrt{a^2+b^2}}\cdot$ Nestes pontos, $f|S^1$ assume, respectivamente, seu valor máximo, igual a $\sqrt{a^2+b^2}$, e seu valor mínimo, igual a $-\sqrt{a^2+b^2}$.

Exemplo 28. Dados $a\in\mathbb{R}^{n+1}$ e uma hiperfície $M\subset\mathbb{R}^{n+1}$, que não contenha o ponto a, determinar o ponto $p\in M$ mais próximo a a. (Se M for um subconjunto fechado de $\mathbb{R}^{n+1}$, sabemos que um tal ponto sempre existe.) Consideremos a função $f(x)=|x-a|$, que é de classe C^∞ em $\mathbb{R}^{n+1}-\{a\}$. Procuramos os pontos onde $f|M$ assume seu valor mínimo. Eles estão entre os pontos críticos de $f|M$, isto é, entre os pontos nos quais $\operatorname{grad} f$ é normal a M. Ora, como $f(x)=\sqrt{\Sigma(x_i-a_i)^2}$, temos $\dfrac{\partial f}{\partial x_i}=\dfrac{x_i-a_i}{|x-a|}$, donde $\operatorname{grad} f(x)=\dfrac{x-a}{|x-a|}\cdot$ Assim, os pontos críticos de $f|M$ (entre os quais se encontram os pontos de M situados a uma distância mínima do ponto a) são os pontos $x\in M$ tais que $x-a$ é normal a M.

Exemplo 29. Seja (a_{ij}) uma matriz real $n\times n$ *simétrica*, isto é, $a_{ij}=a_{ji}$. A ela corresponde uma transformação linear $A\colon\mathbb{R}^n\to\mathbb{R}^n$, definida por $A\cdot x=y$, com $y_i=\sum_j a_{ij}x_j$. A condição $a_{ij}=a_{ji}$ é equivalente a $\langle Ax,y\rangle=\langle x,Ay\rangle$ para quaisquer $x,y\in\mathbb{R}^n$. Diz-se então que a transformação linear A é *auto-adjunta*. Um vetor $x\in\mathbb{R}^n$ chama-se um *autovetor* de A quando $x\neq 0$ e $A\cdot x=\lambda\cdot x$ para algum $\lambda\in\mathbb{R}$. O número λ chama-se então o *autovalor* correspondente ao autovetor x. Em geral, uma transformação linear $A\colon\mathbb{R}^n\to\mathbb{R}^n$ não precisa ter autovetores $x\in\mathbb{R}^n$ nem autovalores reais. (Tome, por exemplo, uma rotação no plano, de ângulo θ, com $0^0<\theta<180°$.) Mostraremos agora que se A é auto-adjunta então existe uma base ortonormal de $\mathbb{R}^n$ formada por autovetores de A. Para isso, introduziremos a forma quadrática $f\colon\mathbb{R}^n\to\mathbb{R}$, dada por $f(x)=\langle A\cdot x,x\rangle$ ou, em termos de coordenadas, $f(x)=\sum_{i,j=1}^n a_{ij}x_ix_j$. Estudaremos os pontos críticos de f na esfera unitária $S^{n-1}\subset\mathbb{R}^n$. Como $\dfrac{\partial f}{\partial x_i}=2\cdot\sum_j a_{ij}x_j$, temos $\operatorname{grad} f(x)=2A\cdot x$. Por outro lado, $S^{n-1}=\varphi^{-1}(1)$, onde $\varphi(x)=\langle x,x\rangle$, donde $\operatorname{grad}\varphi(x)=2\cdot x$. Logo os pontos críticos de $f|S^{n-1}$ são os pontos $u\in S^{n-1}$ tais que $A\cdot u=\lambda\cdot u$. Num tal ponto u, temos $f(u)=\langle Au,u\rangle=\langle\lambda u,u\rangle=\lambda$,

pois $\langle u, u\rangle = 1$. Podemos, portanto, enunciar:

Dada a forma quadrática $f\colon \mathbb{R}^n \to \mathbb{R}$, $f(x) = \langle A\cdot x, x\rangle$, *com* $A\colon \mathbb{R}^n \to \mathbb{R}^n$ *auto-adjunta, um ponto* $u \in S^{n-1}$ *é ponto crítico de* $f|S^{n-1}$ *se, e somente se,* $A\cdot u = \lambda\cdot u$, *onde* $\lambda = f(u)$.

Em particular, se λ_1 é o valor máximo de f no compacto S^{n-1}, atingido no ponto $u_1 \in S^{n-1}$, então λ_1 é o maior autovalor de A. Consideremos agora o subespaço $(n-1)$-dimensional $E = \{x' \in \mathbb{R}^n; \langle x, u_1\rangle = 0\}$, complemento ortogonal do vetor u_1. Se $x \in E$ então $\langle A\cdot x, u_1\rangle = \langle x, A\cdot u_1\rangle = \langle x, \lambda_1 u_1\rangle = \lambda_1\cdot\langle x, u_1\rangle = 0$. Logo $x\in E \Rightarrow A\cdot x \in E$. Por restrição, obtemos uma transformação linear auto-adjunta $A\colon E\to E$. Seja $f(u_2) = \lambda_2$ o valor máximo da forma quadrática f entre os vetores unitários pertencentes a E (isto é, perpendiculares a u_1). Então $A\cdot u_2 = \lambda_2\cdot u_2$. Prosseguindo analogamente, obtemos uma base ortonormal de $\mathbb{R}^n$, $(u_1, u_2, \ldots, u_n)$, formada por autovetores de A.

Exemplo 30. Consideremos a função $f\colon \mathbb{R}^3\to\mathbb{R}$, $f(x,y,z) = x\cdot y\cdot z$, e, para cada $c>0$, procuremos seu valor máximo na superfície $M = \{(x,y,z)\in\mathbb{R}^3; x+y+z = c, y>0, z>0\}$. Esse máximo existe em $\overline{M}$ porque $\overline{M}$ é um conjunto compacto mas, na realidade, está em M pois f é positiva em M e nula em $\overline{M} - M$. Ora $\operatorname{grad} f = (yz, xz, xy)$ e M é uma parte de $\varphi^{-1}(c)$, com $\varphi(x,y,z) = x+y+z$, donde $\operatorname{grad}\varphi = (1,1,1)$. Logo, num ponto de M onde $f|M$ seja máxima, devemos ter $yz = xz = xy = \lambda$ e $x+y+z = c$. Assim, $x = y = z = c/3$ e $f(x,y,z) = c^3/27$. Sendo este valor máximo de f em M, devemos ter $xyz \le \dfrac{c^3}{27} = \left(\dfrac{x+y+z}{3}\right)^3$, sempre que $x>0$, $y>0$, $z>0$ e $x+y+z = c$. Mas c é arbitrário e três números positivos x, y, z têm sempre uma soma c. Podemos então afirmar que, dados 3 números positivos x, y, z, tem-se $xyz \le \left(\dfrac{x+y+z}{3}\right)^3$, ou seja $\sqrt[3]{xyz} \le \dfrac{x+y+z}{3}$; a média geométrica é menor do que ou igual à média aritmética. O mesmo raciocínio, aplicado à função $f\colon\mathbb{R}^n\to\mathbb{R}$, dada por $f(x) = x_1\cdot x_2\cdot\cdots\cdot x_n$, mostra que a média geométrica de n números positivos é menor do que ou igual à média aritmética desses números.

Exemplo 31. Usaremos o método do multiplicador de Lagrange para demonstrar a *desigualdade de Hadamard*: se X é uma matriz $n\times n$ cujas linhas são os vetores $X_i = (x_{i1}, x_{i2}, \ldots, x_{in})$ então $\det . X \le |X_1|\cdot|X_2|\ldots|X_n|$, onde $|X_i|$ é a norma euclidiana de X_i. Isto é evidente se $\det . X = 0$. Caso $\det . X \ne 0$, então todos os vetores-linha são $\ne 0$,

logo $X_i = |X_i| \cdot W_i$, com $|W_i| = 1$, para todo i. Então $\det . X = |X_1| \cdot |X_2| \ldots |X_n| \cdot \det . W$, onde W é a matriz cujas linhas são os vetores unitários $W_1, \ldots, W_n$. A desigualdade de Hadamard ficará provada se mostrarmos que $\det . W \leq 1$. Mais geralmente, mostraremos que se $W = (w_{ij})$ é uma matriz $n \times n$ tal que $\sum_{i,j} w_{ij}^2 = n$ então $\det . W \leq 1$.

Definamos portanto $f, \varphi \colon \mathbb{R}^n \to \mathbb{R}$ pondo $f(X) = \det . X$ e $\varphi(X) = \sum_{i,j} x_{ij}^2$. Temos $\dfrac{\partial \varphi}{\partial x_{ij}}(X) = 2x_{ij}$ e $\dfrac{\partial f}{\partial x_{ij}}(X) = (-1)^{i+j} X_{[ij]}$, onde $X_{[ij]}$ é o determinante da matriz $(n-1) \times (n-1)$, obtida de φ pela omissão da i-ésima linha e da j-ésima coluna. Todo número real $\neq 0$ é um valor regular de φ, logo $M = \varphi^{-1}(n)$ é uma hiperfície (compacta) de classe C^∞ em $\mathbb{R}^{n^2}$. (Esfera de centro 0 e raio $\sqrt{n}$.)

Uma matriz $W = (w_{ij})$ é ponto crítico de $f|M$ se, e somente se, $\sum_{i,j} w_{ij}^2 = n$ e $\operatorname{grad} f(W) = \lambda \cdot \operatorname{grad} \varphi(W)$ para algum λ real. Daí:

$$(*) \qquad (-1)^{i+j}\, W_{[ij]} = 2\lambda \cdot w_{ij} \quad \text{para quaisquer} \quad i, j \in [1, n].$$

Multiplicando por w_{ij}, somando e levando em conta a regra de expansão de um determinante em relação aos elementos de uma linha, temos:

$$n \cdot \det W = \sum_{i,j} (-1)^{i+j}\, w_{ij} W_{[ij]} = 2\lambda \cdot \sum w_{ij}^2 = 2\lambda \cdot n,$$

donde $\det W = 2\lambda$. Agora multipliquemos (*) por w_{ij}, fixemos i e somemos em relação a j. Resulta:

$$\det W = \sum_j (-1)^{i+j}\, w_{ij} W_{[ij]} = 2\lambda \cdot \sum_j w_{ij}^2 = \det W \cdot \sum_j w_{ij}^2 .$$

Suponhamos que W seja um ponto onde $f|M$ assume seu valor máximo. Então $\det W = f(W) \neq 0$ e da igualdade acima vem $|W_i|^2 = \sum_j w_{ij}^2 = 1$ para todo i.

Em seguida, multipliquemos (*) por w_{kj}, com $k \neq i$, e somemos em relação a j. Teremos

$$\sum_j (-1)^{i+j}\, w_{kj} W_{[ij]} = 2\lambda \cdot \sum_j w_{kj} w_{ij} = 2\lambda \cdot \langle W_k, W_i \rangle.$$

Ora, o primeiro somatório acima é zero, por ser o desenvolvimento, em relação aos elementos da i-ésima linha, do determinante de uma matriz

com duas linhas (a i-ésima e a k-ésima) iguais a W_k. Logo $\langle W_k, W_i \rangle = 0$ para $k \neq i$.

Assim, todo ponto $W \in M$ onde $f|M$ atinja o seu valor máximo é uma matriz cujas linhas são vetores unitários, 2 a 2 ortogonais, isto é, W é uma matriz ortogonal. Em particular, $\det W = \pm 1$. Por ser $\det W$ máximo, seu valor é evidentemente 1.

Concluimos que $\det W \leq 1$ para toda $W \in M$, o que demonstra a desigualdade de Hadamard.

Observação: O valor absoluto de $\det. X$ é o volume do paralelepípedo n-dimensional determinado pelos vetores $X_1, \ldots, X_n$, que constituem as linhas de matriz X. A desigualdade de Hadamard significa, geometricamente, que se mantivermos constantes os comprimentos desses vetores, $\det. X$ se torna máximo quando eles forem 2 a 2 perpendiculares, caso em que o volume do paralelepípedo é o produto $|X_1| \cdot |X_2| \ldots |X_n|$ dos comprimentos das suas arestas.

Exercícios

§1.

1.1. Seja $X = \{(x, y) \in \mathbb{R}^2; x > 0 \text{ e } x^2 \leq y \leq 2x^2\}$. Ponha $U = \mathbb{R}^2 - X$. Mostre que, para todo $a \in U$ e todo $v \in \mathbb{R}^2$, existe $\delta > 0$ (dependendo de a e de v) tal que $0 \leq t < \delta \Rightarrow a + tv \in U$. Entretanto U não é um conjunto aberto.

1.2. Seja $U \subset \mathbb{R}^m$ aberto e conexo. Se $f : U \to \mathbb{R}$ possui, em todos os pontos de U, derivadas parciais nulas então f é constante.

1.3. Se $f : U \to \mathbb{R}$, definida no aberto $U \subset \mathbb{R}^m$, assume seu valor máximo (ou mínimo) num ponto $a \in U$ então qualquer derivada parcial de f que exista no ponto a é nula.

1.4. Seja $f : U \to \mathbb{R}$ contínua no aberto limitado $U \subset \mathbb{R}^m$, possuindo derivadas parciais em todos os pontos de U. Se, para todo $a \in \partial U$ tem-se $\lim_{x \to a} f(x) = 0$ então existe $c \in U$ tal que $\frac{\partial f}{\partial x_i}(c) = 0$ para $i = 1, \ldots, m$. ["Teorema de Rolle".]

1.5. Se $f : U \to \mathbb{R}$ possui derivadas parciais, com $\left|\frac{\partial f}{\partial x_i}(x)\right| \leq M$ $(i = 1, \ldots, m)$ em todos os pontos do aberto convexo $U \subset \mathbb{R}^m$ então $|f(x) - f(y)| \leq M|x - y|$ (norma da soma) para quaisquer $x, y \in U$. Conclua que se f possui derivadas parciais limitadas num aberto qualquer ela é contínua (mas não necessariamente uniformemente contínua).

1.6. Seja $A \subset \mathbb{R}^2$ um retângulo aberto, de lados paralelos aos eixos. Se $f : A \to \mathbb{R}$ possui derivadas parciais em todos os pontos de A então, dados (a, b) e

$(a+h, b+k)$ em A existe $\theta \in (0,1)$ tal que $f(a+h,b+k) - f(a,b) = \frac{\partial f}{\partial x}(a+\theta h, b+k) \cdot h + \frac{\partial f}{\partial y}(a, b+\theta k) \cdot k$.

1.7. Se $X \subset \mathbb{R}^m$ é convexo e $Y \subset \mathbb{R}^n$ é arbitrário então $X \times Y \subset \mathbb{R}^{m+n}$ é i-convexo para todo $i = 1, \ldots, m$. Dê exemplo de um aberto conexo $U \subset \mathbb{R}^2$ que seja 1-convexo e 2-convexo mas não seja convexo.

1.8. Seja $f\colon U \to \mathbb{R}$ contínua num aberto $U \subset \mathbb{R}^2$, com derivada parcial $\frac{\partial f}{\partial y}(a,b) > 0$ num ponto $(a,b) \in U$, com $f(a,b) = c$.
(1) Mostre que existe um retângulo $I \times J$, de lados paralelos aos eixos, contido em U, tal que $f(x,y) > c$ quando (x,y) pertence à base superior do retângulo e $f(x,y) < c$ quando (x,y) pertence à base inferior.
(2) Mostre que existe um intervalo I de centro a, no qual está definida uma função $\xi\colon I \to J$ tal que $\xi(a) = b$ e, para todo $x \in I$, $f(x, \xi(x)) = c$. Esta função não é necessariamente única, mas qualquer uma delas é contínua no ponto a.
(3) Se $\frac{\partial f}{\partial y} > 0$ em todos os pontos de U, mostre que ξ é única e é contínua em I.

§2.

2.1. Uma função $f\colon \mathbb{R}^m \to \mathbb{R}$ tal que $f(0) = 0$ e $f(tx) = tf(x)$, para quaisquer $x \in \mathbb{R}^m$ e $t \neq 0$, tem todas as derivadas direcionais na origem, e vale $\frac{\partial f}{\partial v}(0) = f(v)$.

2.2. Seja $f\colon \mathbb{R}^2 \to \mathbb{R}$ definida por $f(x,y) = \frac{x^2 y}{x^2 + y^2}$ se $x^2 + y^2 > 0$ e $f(0,0) = 0$. Para todo caminho $\lambda\colon (-\varepsilon, \varepsilon) \to \mathbb{R}^2$, diferenciável no ponto 0, com $\lambda(0) = (0,0)$, existe a derivada $(f \circ \lambda)'(0)$.

2.3. Sejam $\varphi, \psi\colon \mathbb{R}^2 \to \mathbb{R}$ definidas por

$$\varphi(x,y) = (x^2 - y)^2 \frac{y^2}{x^8}, \psi(x,y) = (x^2 - y)^2 \frac{y^2}{x^7\sqrt{x}} \text{ se } x > 0 \text{ e } 0 < y < x^2.$$

Nos demais pontos de $\mathbb{R}^2$, ponha $\varphi(x,y) = \psi(x,y) = 0$. Mostre que φ e ψ possuem derivadas direcionais $\frac{\partial \varphi}{\partial v}, \frac{\partial \psi}{\partial v}$ em todos os pontos do plano e que essas derivadas dependem linearmente de v. Mostre ainda que ψ é contínua em $\mathbb{R}^2$ mas φ é contínua apenas em $\mathbb{R}^2 - \{0\}$. Finalmente, considerando o caminho diferenciável $\lambda\colon \mathbb{R} \to \mathbb{R}^2$, dado por $\lambda(t) = (t, t^2)$, a função composta $\psi \circ \lambda\colon \mathbb{R} \to \mathbb{R}$ não é derivável no ponto $t = 0$.

2.4. Seja $f\colon \mathbb{R}^m \to \mathbb{R}$ uma função contínua, possuindo todas as derivadas direcionais em qualquer ponto de $\mathbb{R}^m$. Se $\frac{\partial f}{\partial u}(u) > 0$ para todo $u \in S^{m-1}$ então existe um ponto $a \in \mathbb{R}^m$ tal que $\frac{\partial f}{\partial v}(a) = 0$ seja qual for $v \in \mathbb{R}^m$.

§3.

3.1. Seja $f\colon \mathbb{R}^m \to \mathbb{R}$ tal que $f(tx) = |t| f(x)$ para $x \in \mathbb{R}^m$ e $t \in \mathbb{R}$ quaisquer. Se f é diferenciável na origem, então $f(x) = 0$ para todo x.

3.2. Sejam $U \subset \mathbb{R}^m$ um aberto tal que $x \in U$, $\ t > 0 \Rightarrow tx \in U$, e k um número real. Uma função $f\colon U \to \mathbb{R}$ diz-se *positivamente homogênea de grau k* quando $f(tx) = t^k f(x)$ para quaisquer $x \in U$ e $t > 0$. Para todo $k \in \mathbb{R}$ mostre que existe uma função $f\colon \mathbb{R}^m - \{0\} \to \mathbb{R}$, de classe C^∞, positivamente homogênea de grau k, tal que $f(x) > 0$ para todo x e f não é um polinômio.

3.3. Seja $U \subset \mathbb{R}^m$ como no exercício anterior. Se $f\colon U \to \mathbb{R}$ é diferenciável então f é positivamente homogênea de grau k se, e somente se, cumpre a *relação de Euler* $\sum \dfrac{\partial f}{\partial x_i}(x) \cdot x_i = k \cdot f(x)$. Escreva a relação de Euler para a função $f(x) = \langle x, x\rangle^k = |x|^{2k}$.

3.4. Sejam $U \subset \mathbb{R}^m$ aberto e $f\colon U \to \mathbb{R}$ diferenciável no ponto $a \in U$. Prove que existem $\varepsilon > 0$ e $M > 0$ tais que $|h| < \varepsilon \Rightarrow a+h \in U$ e $|f(a+h)-f(a)| \leq M\cdot|h|$.

3.5. Mostre que a seguinte modificação do exercício anterior é falsa: existem $\varepsilon > 0$ e $M > 0$ tais que $|y-a| < \varepsilon$, $|x-a| < \varepsilon \Rightarrow |f(y)-f(x)| \leq M \cdot |y-x|$.

3.6. Seja $f\colon U \to \mathbb{R}$ de classe C^1 no aberto $U \subset \mathbb{R}^m$. Dados $a \in U$ e $\varepsilon > 0$, prove que existe $\delta > 0$ tal que

$$x, y \in U,\ |x-a| < \delta,\ |y-a| < \delta \Rightarrow f(y) - f(x) = f'(a)(y-x) + r(x,y),$$

onde $|r(x,y)| \leq \varepsilon|x-y|$.

3.7. Uma função holomorfa que só assume valores reais num aberto conexo é constante. (Idem, para valores numa reta qualquer do plano.)

3.8. Sejam $f = u + iv$ uma função holomorfa e φ, ψ caminhos diferenciáveis, com valores no domínio de f, tais que $u \circ \varphi$ e $v \circ \psi$ são constantes. Se $\varphi(s) = \psi(t)$ e $f'(\varphi(s)) \neq 0$ então $\langle \varphi'(s), \psi'(t)\rangle = 0$. ("As curvas de nível da parte real e da parte imaginária de uma função holomorfa cortam-se ortogonalmente".)

3.9. Sejam $U \subset \mathbb{R}^m$ aberto, $f\colon U \to \mathbb{R}$ diferenciável no ponto $a \in U$ e $M = \{(x,y) \in \mathbb{R}^{m+1}; x \in U, y = f(x)\}$ o gráfico de f. O conjunto E dos vetores $v = (\alpha_1, \ldots, \alpha_{m+1}) \in \mathbb{R}^{m+1}$ tais que $\alpha_{m+1} = \sum\limits_{i=1}^{m} \alpha_i \cdot \dfrac{\partial f}{\partial x_i}(a)$ é um subespaço vetorial de dimensão m em $\mathbb{R}^{m+1}$. Mostre que E coincide com o conjunto dos vetores velocidade $\lambda'(0)$ dos caminhos $\lambda\colon (-\varepsilon, \varepsilon) \to \mathbb{R}^{m+1}$, diferenciáveis no ponto $t = 0$, com $\lambda(0) = (a, f(a))$ e tais que $\lambda(t) \in M$ para todo t. Determine $\beta_1, \ldots, \beta_{m+1}$ de modo que o vetor $v = (\beta_1, \ldots, \beta_{m+1})$ seja não-nulo e ortogonal ao subespaço E.

3.10. Seja $f\colon I \to \mathbb{R}^n$ um caminho de classe C^2 que representa o movimento de um ponto de massa m, o qual se desloca sob a ação de um campo de forças conservativo. Isto significa que existe uma função $u\colon U \to \mathbb{R}$, de classe C^1, chamada a *energia potencial*, tal que $f(I) \subset U$ e $m \cdot f''(t) = -\operatorname{grad} u(f(t))$ para todo $t \in I$. Defina a *energia cinética* do ponto $f(t)$ como $m|f'(t)|^2/2$. Prove o *princípio da conservação da energia*: em cada instante t, a soma da energia cinética com a energia potencial é igual a uma constante c.

3.11. Na parte (2) do Exercício 1.8 deste capítulo, mostre que se f é diferenciável no ponto (a,b) então qualquer função ξ tal que $\xi(a) = b$ e $f(x, \xi(x)) = c$ para todo $x \in I$ é diferenciável no ponto a.

3.12. Seja $f\colon U \to \mathbb{R}$ diferenciável, positivamente homogênea de grau 1 num aberto $U \subset \mathbb{R}^m$ contendo zero. Mostre que f é a restrição a U de uma transformação linear de $\mathbb{R}^m$ em $\mathbb{R}$. Conclua que a função $f\colon \mathbb{R}^2 \to \mathbb{R}$, dada por $f(x,y) = x^3/(x^2+y^2)$, $f(0,0)=0$, não é diferenciável na origem.

3.13. Seja $f\colon \mathbb{R}^m \to \mathbb{R}$ diferenciável, tal que $f(x/2) = f(x)/2$ para todo $x \in \mathbb{R}^m$. Prove que f é linear.

§4.

4.1. Todo funcional linear $f\colon \mathbb{R}^m \to \mathbb{R}$ é diferenciável e $df(x)\cdot v = f\cdot v$ para quaisquer $x, v \in \mathbb{R}^m$.

4.2. Seja $f\colon U \to \mathbb{R}$ uma função que possui todas as derivadas direcionais $\dfrac{\partial f}{\partial v}(a)$ num ponto $a \in U$, $U \subset \mathbb{R}^m$ aberto. Se não existirem pelo menos $m-1$ vetores v, linearmente independentes, tais que $\dfrac{\partial f}{\partial v}(a) = 0$, então f não é diferenciável no ponto a.

4.3. Dada $f\colon U \to \mathbb{R}$ no aberto $U \subset \mathbb{R}^m$, defina $f^k\colon U \to \mathbb{R}$ pondo $f^k(x) = f(x)^k$. Prove que f^k é diferenciável e que $df^k(x)\cdot v = k\cdot f^{k-1}(x)\cdot df(x)\cdot v$ para $x \in U$ e $v \in \mathbb{R}^m$.

4.4. Para cada uma das funções abaixo, escreva a diferencial sob a forma

$$df(x) = \frac{\partial f}{\partial x_1}(x)\cdot dx_1 + \cdots + \frac{\partial f}{\partial x_m}(x)\cdot dx_m$$

e use esta expressão para calcular $df(x)\cdot v$ para x e v dados.

(a) $f\colon \mathbb{R}\times(\mathbb{R}-\{0\}) \to \mathbb{R}$, $f(x,y) = x/y$. Calcule $df(x,y)\cdot v$ com $v = (tx, ty)$ e relacione este resultado com a curvas de nível de f.

(b) $f\colon \mathbb{R}^3-\{0\} \to \mathbb{R}$, $f(x,y,z) = \left(\sqrt{x^2+y^2+z^2}\right)^{-1}$. Mostre que $df(x,y,z)\cdot v = 0$ se, e somente se, v é perpendicular a (x,y,z). Calcule $df(x,y,z)\cdot v$ para $x=1$, $y=2$, $z=3$ e $v=(4,2,2)$.

(c) $f\colon \mathbb{R}^2 - \{0\} \to \mathbb{R}$, $f(z) = \log|z|$. Calcule $df(z)\cdot v$ com $z = (x,y)$ e $v = (-y, x)$.

4.5. Considere em $\mathbb{R}^m$ a norma euclidiana. Se $f\colon \mathbb{R}^m - \{0\} \to \mathbb{R}$ é definida por $f(x) = |x|^a$, com $a \in \mathbb{R}$, então $df(x)\cdot v = a|x|^{a-2}\langle x, v\rangle$ para todo $v \in \mathbb{R}^m$.

4.6. Para $1 \le i,\ j \le m$ seja $f_{ij}\colon \mathbb{R}^{m^2} \to \mathbb{R}$ definida por $f_{ij}(X) = (i,j)$-ésimo elemento da matriz X^2. Para $X, V \in \mathbb{R}^{m^2}$ quaisquer, mostre que $df_{ij}(X)\cdot V$ é o ij-ésimo elemento da matriz $XV+VX$. Obtenha um resultado análogo com X^3 em vez de X^2. Generalize.

4.7. Seja $f\colon U \to \mathbb{R}$ definida no aberto $U \subset \mathbb{R}^m$. Dado $a \in U$, suponha que, para todo caminho $\lambda\colon (-\varepsilon,\varepsilon) \to U$, com $\lambda(0) = a$, que possua vetor velocidade $v = \lambda'(0)$ no ponto $t=0$, o caminho composto $f\circ\lambda\colon (-\varepsilon,\varepsilon) \to \mathbb{R}$ também possua vetor velocidade $(f\circ\lambda)'(0) = T\cdot v$, onde $T\colon \mathbb{R}^m \to \mathbb{R}$ é linear. Prove que, nestas condições, f é diferenciável no ponto a.

4.8. Seja $f\colon U \to \mathbb{R}$ diferenciável no aberto $U \subset \mathbb{R}^m$. Suponha $df(a) \neq 0$ para um certo $a \in U$ e considere o vetor unitário $u \in \mathbb{R}^m$ tal que $df(a)\cdot u = \max\{df(a)\cdot h; |h| = 1\}$. Se $v \in \mathbb{R}^m$ é tal que $df(a)\cdot v = 0$, mostre que v é perpendicular a u.

4.9. Seja $f\colon \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ dada por $f(x,y) = \langle x, y\rangle$. Mostre que f é diferenciável e que $df(x,y)\cdot(v,w) = \langle v, y\rangle + \langle x, w\rangle$. Generalize, considerando uma forma bilinear $\varphi\colon \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$ qualquer. Generalize ainda mais, tomando $\psi\colon \mathbb{R}^{m_1} \times \cdots \times \mathbb{R}^{m_k} \to \mathbb{R}$ k-linear. Obtenha a diferencial da função determinante como caso particular.

4.10. Prove que $f\colon \mathbb{R}^2 \to \mathbb{R}$ é diferenciável no ponto $c = (a,b)$ se, e somente se, existem funções $\alpha, \beta\colon \mathbb{R}^2 \to \mathbb{R}$ contínuas na origem, tais que, para todo $(h,k) \in \mathbb{R}^2$, se tem $f(a+h, b+k) = f(a,b) + \alpha\cdot h + \beta\cdot k$, onde $\alpha = \alpha(h,k)$ e $\beta = \beta(h,k)$.

4.11. Seja $U \subset \mathbb{R}^m$ aberto. Se a função diferenciável $f\colon U \to \mathbb{R}$ cumpre a condição de Lipschitz $|f(x) - f(y)| \leq c|x - y|$ então $|df(x)\cdot v| \leq c|v|$ para $x \in U$ e $v \in \mathbb{R}^m$.

4.12. Sejam $U = \{x \in \mathbb{R}^m; |x_i| < 1,\ i = 1, \ldots, m\}$ e $f\colon U \to \mathbb{R}$ uma função diferenciável, com $\left|\dfrac{\partial f}{\partial x_i}(x)\right| \leq 3$ para todo $x \in U$. Então $f(U)$ é um intervalo de comprimento $\leq 3m$.

§5.

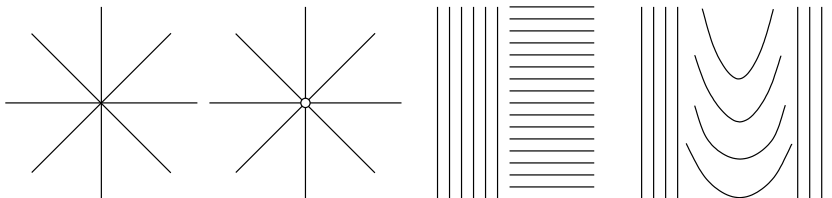
5.1. Dada a transformação linear $A\colon \mathbb{R}^m \to \mathbb{R}^n$, defina as funções $f\colon \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$ e $g\colon \mathbb{R}^m \to \mathbb{R}$ pondo $f(x,y) = \langle A\cdot x, y\rangle$ e $g(x) = \langle A\cdot x, x\rangle$. Determine $\operatorname{grad} f(x,y)$ e $\operatorname{grad} g(x)$.

5.2. Seja $f\colon U \to \mathbb{R}$ diferenciável no aberto $U \subset \mathbb{R}^m$. Dada uma base ortogonal $\{u_1, \ldots, u_m\}$ de $\mathbb{R}^m$, mostre que, para todo $x \in U$, tem-se

$$\operatorname{grad} f(x) = \sum_{i=1}^{m} \frac{1}{|u_i|^2}\,\frac{\partial f}{\partial u_i}(x)\cdot u_i\,.$$

Mais geralmente, dada uma base arbitrária $\{v_1, \ldots, v_m\}$ em $\mathbb{R}^m$, indique com (g^{ij}) a matriz inversa da matriz cujo ij-ésimo elemento é o produto interno $\langle v_i, v_j\rangle$. Mostre que a expressão de $\operatorname{grad} f(x)$ em relação à base $\{v_1, \ldots, v_m\}$ é a seguinte:

$$\operatorname{grad} f(x) = \sum_i \left(\sum_j g^{ij}\,\frac{\partial f}{\partial v_j}\right) v_i\,.$$

5.3. Quais dos esboços abaixo podem representar o conjunto das curvas de nível de uma função $f\colon U \to \mathbb{R}$, de classe C^1 no aberto $U \subset \mathbb{R}^2$?

§6.

6.1. Seja $f\colon [0,b]\times[0,b]\to\mathbb{R}$ contínua. Prove que

$$\int_0^b dx\int_0^x f(x,y)\,dy=\int_0^b dy\int_y^b f(x,y)\,dx$$

e conclua que $\int_0^t dx\int_0^x g(y)\,dy=\int_0^t (t-y)g(y)\,dy$ se $g\colon [0,b]\to\mathbb{R}$ é contínua.

6.2. Seja $f\colon (a,b]\to\mathbb{R}$ contínua. Defina a *integral imprópria* de f pondo $\int_a^b f(x)dx=\lim\limits_{\delta\to 0^+}\int_{a+\delta}^b f(x)\,dx$, quando este limite existir. Defina analogamente as integrais impróprias de funções contínuas $g\colon [a,b)\to\mathbb{R}$ e $h\colon (a,b)\to\mathbb{R}$. Quando a função contínua $f\colon (a,b]\to\mathbb{R}$ é ≥ 0, a integral impropria $\int_a^b f(x)\,dx$ existe se, e somente se, pode-se obter $K>0$ tal que, para todo intervalo fechado $[c,d]\subset(a,b]$, tem-se $\int_c^d f(x)\,dx\leq K$. Para $f\colon (a,b]\to\mathbb{R}$ contínua arbitrária, a integral imprópria $\int_a^b f(x)\,dx$ existe se, e somente se, dado $\varepsilon>0$, pode-se achar $\delta>0$ tal que $a<c<d<a+\delta\Rightarrow|\int_a^d f(x)\,dx|<\varepsilon$. Se $f\colon (a,b]\to\mathbb{R}$ é contínua e $\int_a^b |f(x)|dx$ existe, então $\int_a^b f(x)\,dx$ também existe. (Toda integral absolutamente convergente, é convergente.) Finalmente, se $\int_a^b f(x)\,dx$ existe e $g\colon (a,b]\to\mathbb{R}$ é contínua, com $|g(x)|\leq f(x)$ para todo $x\in(a,b]$, então $\int_a^b g(x)\,dx$ existe.

6.3. Das integrais impróprias abaixo, duas convergem e duas divergem. Quais?

$$\int_{-1}^1\frac{dx}{1-x^2}\,,\quad \int_{-1}^1\frac{dx}{\sqrt{1-x^2}}$$

$$\int_0^1\frac{dx}{x^\alpha}\,,\ \alpha>1,\quad \int_{-1}^1\frac{dx}{\sqrt{(1-x^2)(1-k^2x^2)}}\,,\ k^2<1.$$

6.4. Seja $G\colon (a,b]\times Y\to\mathbb{R}$ contínua, com $Y\subset\mathbb{R}^m$ arbitrário. Diz-se que a integral imprópria $\int_a^b G(t,y)\,dt$ é *uniformemente convergente* em Y quando, para todo $\varepsilon>0$ dado, existir $\delta>0$ tal que $a<c<d<a+\delta\leq b\Rightarrow|\int_c^d G(t,y)\,dt|<\varepsilon$, seja qual for $y\in Y$. Prove que se $F\colon (a,b]\times U\to\mathbb{R}$ (com $U\subset\mathbb{R}^m$ aberto) e suas derivadas parciais em relação às últimas m variáveis são contínuas em $(a,b]\times U$, e existem, para todo $x\in U$, as integrais impróprias

$$\varphi(x)=\int_a^b F(t,x)\,dt\quad\text{e}\quad\psi(x)=\int_a^b\frac{\partial F}{\partial x_i}(t,x)\,dt,$$

sendo a segunda delas uniformemente convergente em cada parte compacta de U, então tem-se $\dfrac{\partial\varphi}{\partial x_i}(x)=\psi(x)$. Noutras palavras, é permitido derivar uma integral imprópria sob o sinal de integração, desde que o resultado seja uma integral uniformemente convergente nas partes compactas. Conclua daí que se

$$\varphi(k)=\int_{-1}^1\frac{dt}{\sqrt{(1-t^2)(1-k^2t^2)}}\,,\ |k|<1\ \text{então}$$

$$\varphi'(k)=\int_{-1}^1\frac{kt^2dt}{\sqrt{(1-t^2)(1-k^2t^2)^3}}\,.$$

6.5. Defina a função de Bessel J_0 pondo

$$J_0(x) = \frac{1}{\pi}\int_{-1}^{1} \frac{\cos xt}{\sqrt{1-t^2}}\, dt.$$

Prove que $J_0'' + (1/x)J_0' + J_0 = 0$ para todo $x > 0$.

6.6. Seja $f\colon (a,b] \times Y \to \mathbb{R}$ contínua, com $Y \subset \mathbb{R}^m$ arbitrário. Se existem constantes c, M, com $0 < c < 1$ e $|f(t,y)| \leq \dfrac{M}{(t-a)^c}$ para quaisquer $t \in (a,b]$ e $y \in Y$, então a integral $\int_a^b f(t,y)\, dt$ converge uniformemente em relação a y.

6.7. Seja $f\colon (a,b] \times [c,d] \to \mathbb{R}$ contínua. Se, para cada $t \in [c,d]$, existe a integral imprópria $\int_a^b f(s,t)\, ds$, a qual converge uniformemente em relação a t, então a integral $\int_a^b f(s,t)\, ds$ é uma função contínua de $t \in [c,d]$ e tem-se

$$\int_c^d dt \int_a^b f(s,t)\, ds = \int_a^b ds \int_c^d f(s,t)\, dt.$$

A igualdade acima significa, em particular, que existe a integral imprópria do segundo membro.

§7.

7.1. Com a notação da Regra da Cadeia, supondo f e g duas vezes diferenciáveis, obtenha uma fórmula para $\dfrac{\partial^2(g \circ f)}{\partial x_i \partial x_j}(a)$.

7.2. Uma função diferenciável $f\colon U \to \mathbb{R}$, definida no aberto $U \subset \mathbb{R}^m$, é de classe C^1 se, e somente se, para cada $h \in \mathbb{R}^m$, a função $\varphi_h\colon U \to \mathbb{R}$, dada por $\varphi_h(x) = df(x) \cdot h$, é contínua. Analogamente, f é duas vezes diferenciável se, e somente se, φ_h é diferenciável.

7.3. Seja $f\colon U \to \mathbb{R}$ duas vezes diferenciável no aberto convexo $U \subset \mathbb{R}^2$. A fim de que $\dfrac{\partial^2 f}{\partial x \partial y}$ seja identicamente nula, é necessário e suficiente que existam funções reais $\varphi\colon I \to \mathbb{R}$, $\psi\colon J \to \mathbb{R}$, duas vezes diferenciáveis em intervalos I, J da reta, tais que $f(x,y) = \varphi(x) + \psi(y)$ para todo $(x,y) \in U$.

7.4. A fim de que uma função duas vezes diferenciável $g\colon \mathbb{R}^2 \to \mathbb{R}$ satisfaça a equação

$$\frac{\partial^2 g}{\partial x^2} = \frac{\partial^2 g}{\partial y^2}$$

é necessário e suficiente que existam funções $\varphi\colon \mathbb{R} \to \mathbb{R}$, $\psi\colon \mathbb{R} \to \mathbb{R}$, duas vezes diferenciáveis, tais que $g(x,y) = \varphi(x+y) + \psi(x-y)$.

7.5. Seja $\mathbb{R}^2 \to \mathbb{R}$ duas vezes diferenciável. Suponha que $\dfrac{\partial^2 f}{\partial y^2} = c^2\, \dfrac{\partial^2 f}{\partial x^2}$ em todos os pontos de $\mathbb{R}^2$, onde c é uma constante. Prove que existem funções $\varphi\colon \mathbb{R} \to \mathbb{R}$, $\psi\colon \mathbb{R} \to \mathbb{R}$, duas vezes diferenciáveis, tais que $f(x,y) = \varphi(x - cy) + \psi(x + cy)$. [Solução da "equação da onda".]

7.6. Seja $U \subset \mathbb{R}^m$ um aberto. Para toda função $f\colon U \to \mathbb{R}$ duas vezes diferenciável, o *Laplaciano* de f é a função $\Delta f\colon U \to \mathbb{R}$, definida por $\Delta f = \dfrac{\partial^2 f}{\partial x_1^2} + \cdots + \dfrac{\partial^2 f}{\partial x_m^2}\cdot$ Prove que se $T\colon \mathbb{R}^m \to \mathbb{R}^m$ é uma transformação linear ortogonal então $\Delta(f \circ T) = (\Delta f) \circ T\colon V \to \mathbb{R}$, onde $V = T^{-1}(U)$. [Invariância do Laplaciano por rotações.]

7.7. Uma função $f\colon U \to \mathbb{R}$, de classe C^2 no aberto $U \subset \mathbb{R}^m$, diz-se *harmônica* quando $\Delta f(x) = 0$ para todo $x \in U$. (A igualdade $\sum\limits_{i=1}^{m} \dfrac{\partial^2 f}{\partial x_i^2} = 0$ chama-se a *equação de Laplace*.) Mostre que as funções abaixo são harmônicas:
(a) $u\colon \mathbb{R}^3 - \{0\} \to \mathbb{R}$, $u(x) = \dfrac{1}{|x|}$, $|x| = \sqrt{\langle x, x\rangle}$.
(b) $p\colon \mathbb{R}^3 \to \mathbb{R}$, $p(x,y,z) = 2x^2 - y^2 - z^2$.
(c) $f\colon \mathbb{R}^2 \to \mathbb{R}$, $f(x,y) = (A \cdot \cos ax + B \cdot \operatorname{sen} ax)e^{ay}$.

§8.

8.1. Seja $f\colon U \to \mathbb{R}$ harmônica no aberto $U \subset \mathbb{R}^2$, isto é, $f \in C^2$ e $\dfrac{\partial^2 f}{\partial x^2} + \dfrac{\partial^2 f}{\partial y^2} = 0$ em todos os pontos de U. Suponha que os pontos críticos de f são todos não-degenerados. Mostre que f não possui máximos nem mínimos locais.

8.2. O conjunto dos pontos em que uma função arbitrária $f\colon X \to \mathbb{R}$, definida num conjunto $X \subset \mathbb{R}^m$, admite um máximo ou um mínimo estrito é enumerável. [Seja $E \subset X$ um subconjunto enumerável denso. Para cada ponto $x \in X$ que seja máximo estrito de f, escolha $e_x \in E$ e $r_x \in \mathbb{Q}$ com $|e_x - x| < r_x/2$ e $y \in B(e_x; r_x) - \{x\} \Rightarrow f(y) < f(x)$. A aplicação $x \mapsto (e_x, r_x)$ é injetiva.]

8.3. Dada $\varphi\colon (a,b) \to \mathbb{R}$ derivável, defina $f\colon (a,b) \times (a,b) \to \mathbb{R}$ pondo $f(x,y) = \int_x^y \varphi(t)\,dt$. Determine os pontos críticos de f, caracterize os pontos críticos não-degenerados, os máximos e mínimos locais e os pontos de sela. Considere $\varphi(t) = 3t^2 - 1$ e esboce as curvas de nível de f neste caso.

8.4. Sejam $f, g\colon U \to \mathbb{R}$ duas vezes diferenciáveis num aberto conexo $U \subset \mathbb{R}^m$. Prove que se $f(a) = g(a)$, $df(a) = dg(a)$ e $d^2f(x) = d^2g(x)$ para todo $x \in U$ então $f = g$.

8.5. Seja $f\colon \mathbb{R}^2 \to \mathbb{R}$ de classe C^∞, com $f(x,0) = f(0,y) = 0$ para quaisquer $x, y \in \mathbb{R}$. Mostre que existe $g\colon \mathbb{R}^2 \to \mathbb{R}$ de classe C^∞ tal que $f(x,y) = g(x,y) \cdot x \cdot y$ para qualquer $(x,y) \in \mathbb{R}^2$.

8.6. Seja $f\colon U \to \mathbb{R}$ de classe C^k $(i \le k \le \infty)$ no aberto convexo $U \subset \mathbb{R}^2$, contendo a origem. Suponha que f e todas as suas derivadas parciais de ordem $\le i$ se anulam na origem. Prove que existem funções $a_0, a_1, \ldots, a_i\colon U \to \mathbb{R}$ de classe C^{k-i}, tais que $f(x,y) = \sum\limits_{j=0}^{i} a_j(x,y) x^j \cdot y^{i-j}$ para todo ponto $(x,y) \in U$.

8.7. Seja $U \subset \mathbb{R}^m$ um aberto convexo. Uma função $f\colon U \to \mathbb{R}$ diz-se *convexa* quando, para $x, y \in U$ e $t \in [0,1]$ quaisquer, tem-se $f((1-t)x + ty) \le (1-t)f(x) + tf(y)$. Seja $E(f) = \{(x,y) \in U \times \mathbb{R}; y \ge f(x)\}$. Mostre que f é convexa se, e somente se, $E(f)$ é um subconjunto convexo de $\mathbb{R}^{m+1}$. (b) Seja

f convexa. Se $x_1, \ldots, x_k \in U$, e $0 \leq t_1, \ldots, t_k \leq 1$, com $\Sigma\, t_i = 1$ então $f(\Sigma\, t_i x_i) \leq \Sigma\, t_k\, f(x_i)$. (c) Se $C \subset \mathbb{R}^m$ é um conjunto convexo então a função $f\colon \mathbb{R}^m \to \mathbb{R}$, dada por $f(x) = \text{dist}(x, C)$, é convexa.

8.8. Seja $U \subset \mathbb{R}^m$ um aberto convexo. Uma função diferenciável $f\colon U \to \mathbb{R}$ é convexa se, e somente se, para $x, x+v \in U$ quaisquer, tem-se $f(x+v) \geq f(x) + df(x) \cdot v$.

8.9. Seja $U \subset \mathbb{R}^m$ aberto e convexo. Uma função duas vezes diferenciável $f\colon U \to \mathbb{R}$ é convexa se, e somente se, para cada $x \in U$, $d^2f(x)$ é uma forma quadrática não-negativa, isto é, $\sum \dfrac{\partial^2 f}{\partial x_i \partial x_j}(x) \cdot \alpha_i \alpha_j \geq 0$ para todo vetor $v = (\alpha_1, \ldots, \alpha_m) \in \mathbb{R}^m$.

8.10. Mostre, diretamente a partir das definições, que toda norma em $\mathbb{R}^m$ é uma função convexa. Se $f\colon \mathbb{R}^m \to \mathbb{R}$ é uma norma proveniente de um produto interno, prove que, para $x \neq 0$ e h qualquer em $\mathbb{R}^m$, tem-se $d^2f(x) \cdot h^2 = (|h|^2|x|^2 - \langle x, h\rangle^2) \cdot |x|^{-3}$ e observe que a convexidade de f é equivalente à desigualdade de Schwarz.

8.11. Seja $f\colon \mathbb{R}^m \to \mathbb{R}$ positivamente homogênea de grau $k \in \mathbb{N}$ e de classe C^k. Então f é um polinômio homogêneo de grau k. Mais precisamente: existe uma forma k-linear φ em $\mathbb{R}^m$ tal que $f(x) = \varphi(x, \ldots, x)$ para todo $x \in \mathbb{R}^m$.

8.12. Por meio de sucessivas mudanças de coordenadas, como foi indicado no Exemplo 18, exprima cada uma as formas quadráticas abaixo como soma de termos do tipo $\pm u^2$ e decida quais são positivas, negativas ou indefinidas

$$A(x,y) = x^2 - 3xy + y^2, \;\; B(x,y,z) = 2xy + yz - 3xz,$$

$$C(x,y,z,t) = x^2 + y^2 + 2xy - xt + 2yt$$

8.13. Seja $f\colon U \to \mathbb{R}$ de classe C^1 no aberto $U \subset \mathbb{R}^2$. Se, para algum ponto $(a,b) \in U$, com $f(a,b) = c$, temos $\dfrac{\partial f}{\partial y}(a,b) > 0$, existe $k > 0$ tal que $\dfrac{\partial f}{\partial y}(x,y) > k$ para todo (x,y) suficientemente próximo de (a,b). Então existe um retângulo $R = [a-\delta, a+\delta] \times [b-\varepsilon, b+\varepsilon] \subset U$ tal que $f(x, b-\varepsilon) < c - k\cdot\varepsilon$ e $f(x, b+\varepsilon) > c + k\cdot\varepsilon$ para todo $x \in [a-\delta, a+\delta]$. Logo $f(R) \supset (c - k\varepsilon, c + k\varepsilon)$. Conclua que se f não possui pontos críticos então, para cada aberto $A \subset U$, $f(A)$ é aberto em $\mathbb{R}$.

8.14. Seja $f\colon \mathbb{R}^m \to \mathbb{R}$ de classe C^1, com $m \geq 2$, tal que para algum $c \in \mathbb{R}$, a imagem inversa $f^{-1}(c)$ é compacta e não-vazia. Mostre que um dos fechados $F = \{x \in \mathbb{R}^m; f(x) \leq c\}$ ou $G = \{x \in \mathbb{R}^m; f(x) \geq c\}$ é compacto. Conclua que f assume um valor máximo ou mínimo em $\mathbb{R}^m$.

8.15. Seja $f\colon U \to \mathbb{R}$ de classe C^2 no aberto $U \subset \mathbb{R}^m$, com $0 \in U$, $f(0) = 0$ e $df(0) = 0$. O gráfico de f é uma hiperfície em $\mathbb{R}^{m+1}$, cuja normal no ponto 0 é o $(m+1)$-ésimo eixo. Seja $p = (0, \ldots, a) \in \mathbb{R}^{m+1}$ um ponto desse eixo, com $a \neq 0$. A função $g\colon U \to \mathbb{R}$, definida por $g(x) = |x|^2 + (f(x) - a)^2 =$ quadrado da distância de p ao ponto $(x, f(x))$ do gráfico de f, tem 0 como ponto crítico, com matriz Hessiana igual a $-2a(H - a^{-1} \cdot I)$, onde H é a matriz Hessiana de f no ponto 0 e I é a matriz identidade $m \times m$. Os pontos $(0, \ldots, 0, x) \in \mathbb{R}^{m+1}$

tais que $1/x$ é autovalor de H chamam-se *pontos focais* (da normal ao gráfico de f no ponto 0). Conclua que a origem é um ponto crítico não-degenerado da função g se, e somente se, p não é um ponto focal.

8.16. O *índice* de um ponto crítico não-degenerado é o número de autovalores negativos da matriz Hessiana. Assim, se o domínio é um aberto de $\mathbb{R}^m$, um ponto de mínimo local tem índice zero, enquanto que um ponto de máximo local tem índice m. No problema anterior, mostre que se p não é um ponto focal, então o índice da origem como ponto crítico da função g é igual ao número de pontos focais situados entre p e 0, contados com suas multiplicidades.

§9.

9.1. Sejam $f\colon \mathbb{R}^2 \to \mathbb{R}$ de classe C^1, com $\dfrac{\partial f}{\partial y} \neq 0$ em todos os pontos, e $\xi\colon I \to \mathbb{R}$ tal que $f(x,\xi(x)) = 0$ para todo $x \in I$. Prove que ξ é de classe C^1.

9.2. Seja $f\colon U \to \mathbb{R}$ contínua no aberto $U \subset \mathbb{R}^2$, tal que $(x^2+y^4)f(x,y)+f(x,y)^3 = 1$ para qualquer $(x,y) \in U$. Prove que $f \in C^\infty$.

9.3. Sejam $f,g\colon \mathbb{R}^n \to \mathbb{R}$ tais que $g(x) = f(x) + (f(x))^5$. Se f é contínua e $g \in C^r$ então $f \in C^r$.

9.4. Seja $f\colon U \to \mathbb{R}$ contínua no aberto $U \subset \mathbb{R}^m$. Se a função $g\colon U \to \mathbb{R}$, dada pela expressão $g(x) = \int_0^{f(x)} (t^2+1)\,dt$, for de classe C^∞, então f também será C^∞.

9.5. No conjunto aberto $U = \{(x,y,z) \in \mathbb{R}^3; x > 0\}$, defina a função $f\colon U \to \mathbb{R}$, pondo $f(x,y,z) = x \operatorname{sen} z - y \cos z$. Mostre que, dado $p = (x,y,z) \in U$, se $f(p) = 0$ então $\dfrac{\partial f}{\partial z}(p) \neq 0$. Seja $V = \{(x,y) \in \mathbb{R}^2; x > 0\}$. Mostre que existe uma infinidade enumerável de funções contínuas $\alpha\colon V \to \mathbb{R}$ tais que $f(x,y,\alpha(x,y)) = 0$ para todo $(x,y) \in V$. Cada uma delas é C^∞, duas quaisquer diferem por uma constante e se for fixado um valor $\alpha(x_0,y_0) = z_0$, então a equação $f(x,y,d(x,y)) = 0$ determina α univocamente. Mostre ainda que $d\alpha(x,y) = \dfrac{xdy - ydx}{x^2+y^2}$.

9.6. No enunciado do Teorema da Função Implícita, suponha apenas que $f\colon U \to \mathbb{R}$ é contínua no aberto $U \subset \mathbb{R}^{n+1}$ e diferenciável no ponto (x_0,y_0), com $f(x_0,y_0) = c$ e $\dfrac{\partial f}{\partial y}(x_0,y_0) \neq 0$. Prove que existem uma bola $B = B(x_0;r) \subset \mathbb{R}^n$, um intervalo $J \subset R$ e pelo menos uma função $\xi\colon B \to J$ tal que $\xi(x_0) = y_0$ e $f(x,\xi(x)) = c$ para todo $x \in B$. Toda ξ nessas condições é diferenciável no ponto x_0. Se f for diferenciável, com $\dfrac{\partial f}{\partial y} \neq 0$ em todos os pontos de U, então B e J podem ser escolhidos de modo que ξ seja única e diferenciável em todos os pontos de B.

9.7. O conjunto $M \subset \mathbb{R}^{m^2}$ das matrizes $m \times m$ de posto $m-1$ é uma hiperfície. Seja P a matriz $m \times m$ cujos elementos são todos nulos, exceto os $m-1$ primeiros da diagonal, que são iguais a 1. Determine T_pM.

9.8. Sejam $M \subset \mathbb{R}^m$ uma hiperfície diferenciável, $p \in M$ e $A_p = p + T_pM = \{p+v; v \in T_pM\}$ a variedade afim tangente a M no ponto p. Defina $\varphi\colon M \to \mathbb{R}$ pondo $\varphi(x) = \dfrac{d(x,A_p)}{|x-p|}$. Mostre que $\lim\limits_{x\to p} \varphi(x) = 0$.

9.9. Mostre que 1 é um valor regular da função $f\colon \mathbb{R}^3 \to \mathbb{R}$, definida por $f(x,y,z) = z^2 + \left(\sqrt{x^2+y^2} - 2\right)^2$. Conclua que o toro gerado pela rotação, em torno do eixo z, de um círculo vertical de raio 1 e centro num ponto do plano $z = 0$ distando 2 da origem, é uma superfície de classe C^∞.

9.10. Seja $f\colon [0,2] \to \mathbb{R}$ contínua, positiva, tal que $\int_0^1 f(x)\,dx = \int_1^2 f(x)\,dx = 1$. Para cada $x \in [0,1]$, prove que existe um único $g(x) \in [1,2]$ tal que $\int_x^{g(x)} f(t)\,dt = 1$. Mostre que a função $g\colon [0,1] \to \mathbb{R}$ assim definida é de classe C^1.

§10.

10.1 Dentre os pontos do elipsóide $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z2}{c^2} = 1$, determine os mais próximos da origem em $\mathbb{R}^3$.

10.2 Determine os pontos críticos da função $f\colon \mathbb{R}^{2m} \to \mathbb{R}$, $f(x,y) = \langle x,y\rangle$, restrita à esfera unitária $|x|^2 + |y|^2 = 1$ e mostre como daí se obtém a desigualdade de Schwarz.

10.3 Sejam $x \geq 0$, $y \geq 0$ e $p > 0$, $q > 0$ com $1/p + 1/q = 1$. Mostre que $xy \leq 1/p\, x^p + 1/q\, y^q$. [Sugestão: não há perda de generalidade em supor $xy = 1$. Determine o mínimo do segundo membro sobre a hipérbole $xy = 1$.] Dado $v = (v_1, \ldots, v_m) \in \mathbb{R}^m$, ponha $|v|_p = \left(\sum_i |v_i|^p\right)^{1/p}$. Use a desigualdade acima para provar que, se $1/p + 1/q = 1$ então, dados $u, v \in \mathbb{R}^m$ vale a *desigualdade de Hölder* $|\langle u,v\rangle| \leq |u|_p \cdot |v|_q$.

Capítulo IV

Integrais Curvilíneas

1 Formas diferenciais de grau 1

Uma *forma diferencial de grau* 1 ou, simplesmente, uma 1-*forma* num subconjunto $X \subset \mathbb{R}^n$ é uma aplicação $\omega\colon X \to (\mathbb{R}^n)^*$, que associa a cada ponto $x \in X$ um funcional linear $\omega(x)$ sobre $\mathbb{R}^n$.

Como sabemos, o espaço dual $(\mathbb{R}^n)^*$ possui uma base canônica $(dx_1, \ldots, dx_n)$, formada pelos funcionais definidos por $dx_i \cdot v = \alpha_i$ se $v = (\alpha_1, \ldots, \alpha_n)$; todo funcional linear sobre $\mathbb{R}^n$ se exprime, de modo único, como combinação linear $a_1 dx_1 + \cdots + a_n dx_n$. Assim, dar uma 1-forma ω num subconjunto $X \subset \mathbb{R}^n$ equivale a definir n funções reais $a_1, \ldots, a_n\colon X \to \mathbb{R}$ tais que, para cada $x \in X$, se tenha:

$$\omega(x) = a_1(x)dx_1 + \cdots + a_n(x)dx_n\,.$$

Escrevemos, neste caso, $\omega = \Sigma\, a_i dx_i$.

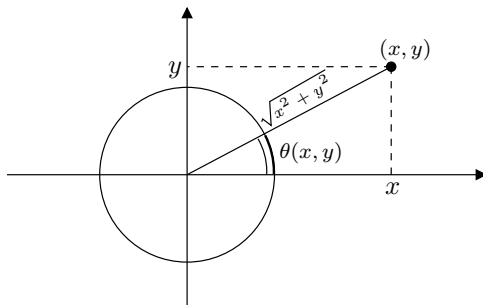
Para cada $x \in X$, a aplicação do funcional $\omega(x)$ no vetor $v = (\alpha_1, \ldots, \alpha_n)$ dá como resultado o número real

$$(*) \qquad \omega(x) \cdot v = a_1(x) \cdot \alpha_1 + \cdots + a_n(x) \cdot \alpha_n\,.$$

Seja $U \subset \mathbb{R}^n$ aberto. Diremos que a 1-forma $\omega\colon U \to (\mathbb{R}^n)^*$ é *de classe* C^k $(0 \leq k \leq \infty)$ quando, para todo vetor $v \in \mathbb{R}^n$, a função $x \mapsto \omega(x) \cdot v$ é de classe C^k em U. Isto ocorre se, e somente se, $\omega = \Sigma\, a_i dx_i$, onde as funções $a_1, \ldots, a_n\colon U \to \mathbb{R}$ são de classe C^k. De fato, esta condição é suficiente em virtude da igualdade (*) acima e é necessária porque $a_i(x) = \omega(x) \cdot e_i$ para todo $x \in U$ e todo $i = 1, \ldots, n$.

O primeiro exemplo de 1-forma de classe C^k é fornecido pela diferencial $df = \Sigma \dfrac{\partial f}{\partial x_i} dx_i$ de uma função $f\colon U \to \mathbb{R}$, de classe C^{k+1}. Para cada $x \in U$ e cada $v = (\alpha_1, \ldots, \alpha_n) \in \mathbb{R}^n$, temos

$$df(x) \cdot v = \Sigma \frac{\partial f}{\partial x_i}(x) \cdot \alpha_i \,.$$

Outro exemplo importante é a 1-forma "elemento de ângulo", definida em $U = \mathbb{R}^2 - \{0\}$, que apresentaremos a seguir. Ela provém da tentativa de definir uma função $\theta\colon U \to \mathbb{R}$ tal que, em cada ponto $(x, y) \in U$, $\theta(x, y)$ seja uma determinação, em radianos, do ângulo que o eixo das abcissas forma com o raio que liga a origem ao ponto (x, y).

Consideremos inicialmente um aberto qualquer $V \subset \mathbb{R}^2 - \{0\}$.

Uma função $\theta\colon V \to \mathbb{R}$ chama-se uma *função-ângulo* quando, em cada ponto $(x, y) \in V$, o seu valor $\theta = \theta(x, y)$ cumpre as seguintes condições:

$$\cos\theta = \frac{x}{\sqrt{x^2 + y^2}}\,,\ \operatorname{sen}\theta = \frac{y}{\sqrt{x^2 + y^2}}\,.$$

Daí resulta, naturalmente, $\operatorname{tg}\theta = y/x$ se $x \neq 0$ e $\operatorname{cotg}\theta = x/y$ se $y \neq 0$.

Ora, não existe uma função-ângulo *contínua* $\theta\colon U \to \mathbb{R}$, definida em todo o aberto $U = \mathbb{R}^2 - \{0\}$. Na verdade, basta que um aberto V contenha um círculo C, com centro na origem, para que não exista uma função-ângulo contínua $\theta\colon V \to \mathbb{R}$. Com efeito, sendo r o raio do círculo C, para cada ponto $(x, y) \in C$ teríamos $(x, y) = (r \cdot \cos\theta(x, y), r \cdot \operatorname{sen}\theta(x, y))$. Daí resultaria que a restrição $\theta\colon C \to \mathbb{R}$ seria injetiva, o que é absurdo. (Nenhuma função real f, contínua num círculo C, pode ser injetiva pois se $f(p) = a$ e $f(q) = b$ são respectivamente o valor mínimo e o valor máximo de f no compacto C, os dois arcos de círculo com extremidades p e q se aplicam ambos por f sobre o intervalo $[a, b]$.)

Mas, para toda semi-reta fechada $\rho \subset \mathbb{R}^2$, partindo da origem, pondo $V_\rho = \mathbb{R}^2 - \rho$, mostraremos que existe uma função-ângulo contínua $\theta\colon V_\rho \to \mathbb{R}$. Com efeito, seja $p = (\cos a, \operatorname{sen} a)$ o ponto onde a semi-reta ρ intersecta o círculo unitário S^1. Sabemos (veja Exemplo G, §8, Capítulo I) que $\xi\colon (a, a+2\pi) \to S^1 - \{p\}$, definida por $\xi(t) = (\cos t, \operatorname{sen} t)$, é um homeomorfismo. Então definimos uma função-ângulo $\theta\colon V_\rho \to \mathbb{R}$ pondo $\theta(z) = \xi^{-1}\left(\frac{z}{|z|}\right)$ para todo $z = (x, y) \in \mathbb{R}^2 - \rho$.

Podemos então concluir que, embora não exista uma função-ângulo contínua $\theta\colon \mathbb{R}^2 - \{0\} \to \mathbb{R}$, definida globalmente em $\mathbb{R}^2 - \{0\}$, ela existe localmente, isto é, todo ponto $z \in \mathbb{R}^2 - \{0\}$ pertence a um aberto no qual se pode definir uma função-ângulo contínua.

Convém observar que toda função-ângulo contínua $\theta\colon V \to \mathbb{R}$ é, de fato, uma função C^∞.

Mais geralmente, se $f\colon V \to \mathbb{R}$ é contínua no aberto $V \subset \mathbb{R}^n$ e as funções $\alpha, \beta\colon V \to \mathbb{R}$ definidas por $\alpha(x) = \cos f(x)$, $\beta(x) = \operatorname{sen} f(x)$, são diferenciáveis (respect. de classe C^k) então f é diferenciável (respect. de classe C^k). Com efeito, para todo $x \in V$, pelo menos uma das funções cos ou sen tem derivada $\neq 0$ no ponto $f(x)$. Então, digamos, cos é um difeomorfismo C^∞ de um intervalo aberto I, de centro $f(x)$, sobre um intervalo aberto J. Seja $\varphi\colon J \to I$ o inverso de cos. Como f é contínua, existe um disco aberto W, de centro x, tal que $f(W) \subset I$. Em W, temos $\varphi \circ (\cos \circ f) = f$. Como $\alpha = \cos \circ f$ é diferenciável (respect. de classe C^k) e $\varphi \in C^\infty$, segue-se que f é diferenciável (respect. de classe C^k). No caso presente, se $\theta\colon V \to \mathbb{R}$ for uma função-ângulo contínua, teremos $\cos\theta(x, y) = x/\sqrt{x^2+y^2}$ e $\operatorname{sen}\theta(x, y) = y/\sqrt{x^2+y^2}$, ambas funções de classe C^∞ em V. Então $\theta \in C^\infty$.

Resulta imediatamente da definição que se $\theta\colon V \to \mathbb{R}$ é uma função-ângulo e $\varphi\colon V \to \mathbb{R}$ é uma função contínua tal que, para cada $z \in V$, existe um inteiro $k = k_z$, com $\theta(z) - \varphi(z) = 2k\pi$, então φ também é uma função-ângulo.

Vale a seguinte recíproca:

Duas funções-ângulo contínuas $\theta_1, \theta_2\colon V \to \mathbb{R}$, definidas num aberto *conexo* $V \subset \mathbb{R}^2 - \{0\}$, diferem por um múltiplo inteiro, constante, de 2π. Com efeito, como $\theta_1(x, y)$ e $\theta_2(x, y)$ têm o mesmo seno e o mesmo cosseno, existe, para cada $(x, y) \in V$ um inteiro $k(x, y)$ tal que $\theta_2(x, y) = \theta_1(x, y) + 2k(x, y)\pi$. Sendo V conexo, a função contínua inteira $k(x, y) = \frac{1}{2\pi}\,[\theta_2(x, y) - \theta_1(x, y)]$ é constante.

Podemos resumir a discussão acima com o

Teorema 1. *Seja* $U = \mathbb{R}^2 - \{0\}$. *Então:*

1º) *Não existe função-ângulo contínua* $\theta\colon U \to \mathbb{R}$*;*

2º) *Todo ponto* $z \in U$ *pertence a um conjunto aberto* V *no qual se pode definir uma função-ângulo contínua* $\theta\colon V \to \mathbb{R}$*;*

3º) *Toda função-ângulo contínua é de classe* C^∞*;*

4º) *Se* V *é conexo e* $\theta_1, \theta_2\colon V \to \mathbb{R}$ *são funções-ângulo contínuas, existe* $k \in \mathbb{Z}$ *tal que* $\theta_2(z) - \theta_1(z) = 2k\pi$ *para todo* $z \in V$*.*

Em particular, se θ_1 , θ_2 são funções-ângulo contínuas, definidas num aberto conexo V, temos $d\theta_1 = d\theta_2$. Assim, embora não exista $\theta\colon U \to \mathbb{R}$, é como se existisse a diferencial $d\theta\colon U \to (\mathbb{R}^2)^*$! Como $d\theta$ não é, na realidade, a diferencial de uma função globalmente definida em $U = \mathbb{R}^2 - \{0\}$, usaremos o símbolo $\delta\theta$ para representá-la.

Mais precisamente, definiremos a 1-forma $\delta\theta\colon U \to (\mathbb{R}^2)^*$, de classe C^∞ em $U = \mathbb{R}^2 - \{0\}$, a qual chamaremos *elemento de ângulo*, pondo, para cada $(x, y) \in U$,

$$\delta\theta(x, y) = \frac{-y}{x^2 + y^2}\,dx + \frac{x}{x^2 + y^2}\,dy.$$

Teorema 2. *Se existe uma função-ângulo contínua (e portanto,* C^∞*)* $\theta\colon V \to \mathbb{R}$*, definida no aberto* $V \subset \mathbb{R}^2 - \{0\}$*, então sua diferencial coincide, em* V*, com a forma* $\delta\theta$*.*

Reciprocamente, se a forma $\delta\theta$ *é, no aberto conexo* $V \subset \mathbb{R}^2 - \{0\}$*, a diferencial de alguma função* $f\colon V \to \mathbb{R}$ *então, para uma constante conveniente c, a função* $\theta(x, y) = f(x, y) - c$ *é uma função-ângulo contínua em* V*.*

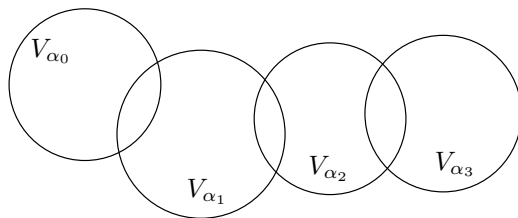
Demonstração: Para cada $z{=}(x, y){\in} V$, temos $\operatorname{sen}\theta(z){=}y/\sqrt{x^2{+}y^2}$. Derivando ambos os membros desta igualdade em relação a x, vem:

$$\cos\theta(z)\cdot\frac{\partial\theta}{\partial x}(z) = \frac{-xy}{(x^2 + y^2)\sqrt{x^2 + y^2}}\,, \text{ donde}$$

$$\frac{\partial\theta}{\partial x}(z) = \frac{-y}{x^2 + y^2}\,, \text{ se } x \neq 0.$$

De modo análogo, derivando a igualdade $\cos\theta(z) = x/\sqrt{x^2+y^2}$ em relação a y, obtemos $\partial\theta/\partial y = x/(x^2+y^2)$ nos pontos em que $y \neq 0$. Daí concluimos que

$$d\theta(x,y) = \frac{-y}{x^2+y^2}\,dx + \frac{x}{x^2+y^2}\,dy$$

nos pontos $(x,y) \in V$ tais que $x \neq 0$ e $y \neq 0$. Como $d\theta$ é contínua e esses pontos formam um conjunto denso em V, por passagem ao limite, segue-se que $d\theta$ tem a expressão acima, ou seja, coincide com $\delta\theta$ em todos os pontos de V.

Para demonstrar a recíproca, escrevamos $V = \cup V_\alpha$, onde cada V_α é uma bola aberta, na qual está definida uma função-ângulo contínua $\alpha\colon V_\alpha \to \mathbb{R}$. Em qualquer bola V_α, como acabamos de mostrar, temos $d\alpha = \delta\theta = df$, logo $f - \alpha$ é constante em V_α. Escolhamos um índice α_0; seja $c \in \mathbb{R}$ tal que $f(z) - \alpha_0(z) = c$ para todo $z \in V_{\alpha_0}$. Afirmamos que a função $\theta\colon V \to \mathbb{R}$, definida por $\theta(z) = f(z) - c$, é uma função-ângulo. Para provar isto, mostraremos que, para todo índice α, existe um inteiro k_α tal que $\theta(z) - \alpha(z) = 2k_\alpha\pi$ seja qual for $z \in V_\alpha$. Com efeito, se $V_{\alpha_1} \cap V_{\alpha_2} \neq \varnothing$, existe $k_{12} \in \mathbb{Z}$ com $\alpha_1(z) - \alpha_2(z) = 2k_{12}\pi$ para todo $z \in V_{\alpha_1} \cap V_{\alpha_2}$. Dado α arbitrário, como V é conexo, existem $\alpha_1, \ldots, \alpha_p = \alpha$ tais que $V_{\alpha_{j-1}} \cap V_{\alpha_j} \neq \varnothing$ para $j = 1, \ldots, p$. Então $\theta(z) - \alpha_1(z) = \alpha_0(z) - \alpha_1(z) = 2k_{01}\pi$ se $z \in V_{\alpha_0} \cap V_{\alpha_1}$. Como $\theta - \alpha_1$ é constante em V_{α_1}, temos $\theta(z) - \alpha_1(z) = 2k_{01}\pi$ para todo $z \in V_{\alpha_1}$. Analogamente, para todo $z \in V_{\alpha_1} \cap V_{\alpha_2}$, temos $\theta(z) - \alpha_2(z) = \alpha_1(z) - \alpha_2(z) + 2k_{01}\pi = 2k_{12}\pi + 2k_{01}\pi = 2(k_{01} + k_{12})\pi$. Novamente, como $\theta - \alpha_2$ é constante em V_{α_2}, segue-se que $\theta(z) - \alpha_2(z) = 2(k_{01} + k_{12})\pi$ para todo $z \in V_{\alpha_2}$. E assim por diante: chegaremos a $\theta(z) - \alpha(z) = \theta(z) - \alpha_p(z) = 2k_\alpha\pi$, para todo $z \in V_\alpha$, onde $k_\alpha = k_{01} + k_{12} + \cdots + k_{p-1,p}$. Ora, todo ponto $z = (x,y) \in V$ pertence a algum V_α. Então $\cos\theta(x,y) = \cos(\alpha(x,y) + 2k_\alpha\pi) = \cos\alpha(x,y) = x/\sqrt{x^2+y^2}$ e, analogamente, $\operatorname{sen}\theta(x,y) = y/\sqrt{x^2+y^2}$, como queríamos demonstrar.

Observação: Se o aberto $V \subset \mathbb{R}^2 - \{0\}$ não é conexo mas existe uma função $f\colon V \to \mathbb{R}$ cuja diferencial coincide com a forma $\delta\theta$, ainda assim existe uma função-ângulo $\theta\colon V \to \mathbb{R}$. A função θ é definida, em cada componente conexa de V separadamente, subtraindo de f uma constante, (que não é a mesma em componentes distintas) conforme o teorema acima.

O corolário seguinte resume o essencial do Teorema 2.

Corolário. *A fim de que exista uma função-ângulo contínua, definida no aberto $V \subset \mathbb{R}^2 - \{0\}$, é necessário e suficiente que a forma elemento de ângulo seja, em V, a diferencial de uma função.*

2 Integral de Stieltjes

A integral de uma 1-forma ao longo de um caminho, que definiremos no parágrafo seguinte, é uma aplicação da Integral de Stieltjes, da qual faremos agora uma rápida exposição.

Sejam $f, \alpha\colon [a,b] \to \mathbb{R}$ duas funções reais definidas no intervalo compacto $[a,b]$. A cada partição pontilhada $P^* = (P, \xi)$, $P = \{a = t_0 < t_1 < \cdots < t_k = b\}$, $t_{i-1} \leq \xi_i \leq t_i$, do intervalo $[a,b]$, associaremos a *soma de Stieltjes* $\Sigma(f, \alpha, P^*) = \Sigma(P^*)$, definida por

$$\Sigma(P^*) = \sum_{i=1}^{k} f(\xi_i)[\alpha(t_i) - \alpha(t_{i-1})].$$

Definiremos a *integral de Stieltjes* de f relativamente a α pondo

$$\int_a^b f(t)\, d\alpha = \lim_{|P| \to 0} \Sigma(P^*).$$

Bem entendido, o limite acima pode não existir. Caso exista, f diz-se *integrável* à Stieltjes, relativamente a α.

Lembremos o significado do limite.

Em primeiro lugar, $|P|$ = maior dos comprimentos $t_i - t_{i-1}$ chama-se a *norma* da partição P. A igualdade acima significa que, dado arbitrariamente $\varepsilon > 0$, existe $\delta > 0$ tal que $|\int_a^b f(t)\, d\alpha - \Sigma(P^*)| < \varepsilon$, seja qual for a partição pontilhada P^*, com $|P| < \delta$.

Vale o

Critério de Cauchy. *A fim de que exista* $\lim_{|P|\to 0} \Sigma(P^*)$ *é necessário e suficiente que, para todo* $\varepsilon > 0$ *dado, se possa obter* $\delta > 0$ *tal que* $|\Sigma(P^*) - \Sigma(Q^*)| < \varepsilon$, *sejam quais forem as partições pontilhadas* P^* *e* Q^*, *com* $|P| < \delta$ *e* $|Q| < \delta$.

Demonstração: A necessidade é evidente. Para provar a suficiência, observemos que, sendo válida a condição, para toda seqüência de partições pontilhadas P_n^*, com $\lim_{n\to\infty} |P_n| = 0$, a seqüência $(\Sigma(P_n^*))$ é de Cauchy, logo converge para um limite $I \in \mathbb{R}$. Este limite I não depende da seqüência (P_n^*) pois se tivéssemos $\lim_{n\to\infty} \Sigma\, Q_n^* = J$, com $\lim |Q_n| = 0$ e $J \neq I$, então a seqüência $(P_1^*, Q_1^*, P_2^*, Q_2^*, \dots)$ originaria uma seqüência divergente de somas de Stieltjes. Vê-se imediatamente que este número I é o limite procurado.

Valem, para a integral de Stieltjes, as regras operatórias usuais, como:

$$\int_a^b (f+g)\, d\alpha = \int_a^b f\, d\alpha + \int_a^b g\, d\alpha;$$

$$\int_a^b (c \cdot f)\, d\alpha = c \cdot \int_a^b f\, d\alpha;$$

$$\int_a^c f\, d\alpha + \int_c^b f\, d\alpha = \int_a^b f\, d\alpha.$$

Na última igualdade acima, devemos supor, que $\int_a^b f\, d\alpha$ existe. Então existem as integrais $\int_a^c f\, d\alpha$ e $\int_c^b f\, d\alpha$, para todo $c \in (a, b)$, e vale a igualdade. A este respeito, veja o Exercício 2.3 adiante.

Se tomarmos $\alpha(t) = t$ (função identidade), a integral de Stieltjes $\int_a^b f(t)\, d\alpha$ reduz-se à integral de Riemann $\int_a^b f(t)\, dt$.

Lembremos que uma função $\alpha\colon [a, b] \to \mathbb{R}$ diz-se de *variação limitada* quando α é um caminho retificável no espaço $\mathbb{R}^1$. No caso afirmativo, o comprimento de α chamar-se-á a *variação total* de α, indicada pela notação $V(\alpha)$, em vez de $\ell(\alpha)$. Assim, $V(\alpha) = \sup_P \Sigma|\alpha(t_i) - \alpha(t_{i-1})|$, o supremo sendo estendido a todas as partições P do intervalo $[a, b]$.

O teorema abaixo estabelece a condição suficiente usual para assegurar a existência da integral de Stieltjes.

Teorema 3. *Se* f *é contínua e* α *tem variação total limitada então existe* $\int_a^b f(t)\, d\alpha$.

Demonstração: Usaremos o Critério de Cauchy para provar que existe $\lim_{|P|\to 0} \Sigma(P^*)$. Seja, então, dado $\varepsilon > 0$. Pela continuidade uniforme de f, existe $\delta > 0$ tal que, para toda partição P de $[a, b]$ com $|P| < \delta$, a oscilação de f em cada subintervalo de P é inferior a $\varepsilon/2V(\alpha)$. Sejam agora P^* e Q^* duas partições pontilhadas quaisquer, com $|P| < \delta$ e $|Q| < \delta$. Temos $P^* = (P, \xi)$ e $P = \{t_0 < t_1 < \cdots < t_n\}$. Se $(P \cup Q)^* = (P \cup Q, \zeta)$ é qualquer pontilhamento de $P \cup Q = \{s_0 < s_1 < \cdots < s_m\}$, escrevamos $j \subset i$ para indicar $[s_{j-1}, s_j] \subset [t_{i-1}, t_i]$. Então para cada $i = 1, \ldots, n$, temos $\alpha(t_i) - \alpha(t_{i-1}) = \sum_{j\subset i}[\alpha(s_j) - \alpha(s_{j-1})]$, donde $f(\xi_i)[\alpha(t_i) - \alpha(t_{i-1})] = \sum_{j\subset i} f(\xi_i)[\alpha(s_j) - \alpha(s_{j-1})]$. Logo

$$\Sigma(P^*) = \sum_{i=1}^{n} \left\{ \sum_{j\subset i} f(\xi_i)[\alpha(s_j) - \alpha(s_{j-1})] \right\}$$

e $\Sigma(P \cup Q)^* = \sum_{j=1}^{m} f(\zeta_j)[\alpha(s_j) - \alpha(s_{j-1})]$. Portanto:

$$\Sigma(P^*) - \Sigma(P \cup Q)^* = \sum_i \left\{ \sum_{j\subset i} [f(\xi_i) - f(\zeta_j)][\alpha(s_j) - \alpha(s_{j-1})] \right\}.$$

Mas, como $|P| < \delta$, $j \subset i \Rightarrow |f(\xi_i) - f(\zeta_j)| < \varepsilon/2 \cdot V(\alpha)$. Assim

$$|\Sigma(P^*) - \Sigma(P \cup Q)^*| < \frac{\varepsilon}{2 \cdot V(\alpha)} \sum_j |\alpha(s_j) - \alpha(s_{j-1})| \leq \frac{\varepsilon}{2} \cdot$$

Analogamente, $|Q| < \delta \Rightarrow |\Sigma(Q^*) - \Sigma(P \cup Q)^*| < \varepsilon/2$. Logo $|P| < \delta$, $|Q| < \delta \Rightarrow |\Sigma(P^*) - \Sigma(Q^*)| < \varepsilon$. Cumpre-se o Critério de Cauchy, portanto existe $\lim_{|P|\to 0} \Sigma(P^*) = \int_a^b f(t)\, d\alpha$.

Em seguida mostraremos que, nos casos mais comuns, a integral de Stieltjes se reduz a uma integral de Riemann, substituindo-se $d\alpha$ por $\alpha'(t)dt$.

Teorema 4. *Se f é contínua e α é de classe C^1 em $[a, b]$ então $\int_a^b f(t)\, d\alpha = \int_a^b f(t)\alpha'(t)\, dt$.*

Demonstração: Para toda partição pontilhada $P^* = (P, \xi)$ temos, em virtude do Teorema do Valor Médio:

$$\Sigma(P^*) = \sum_i f(\xi_i)[\alpha(t_i) - \alpha(t_{i-1})] = \sum_i f(\xi_i)\alpha'(\eta_i)(t_i - t_{i-1}),$$

onde $t_{i-1} < \eta_i < t_i$ para todo i. Logo

$$\Sigma(P^*) = \sum_i f(\xi_i)\alpha'(\xi_i)(t_i - t_{i-1}) + \\ + \sum_i f(\xi_i)[\alpha'(\eta_i) - \alpha'(\xi_i)](t_i - t_{i-1}).$$

Seja $M = \sup\{|f(t)|; a \leq t \leq b\}$. Pela continuidade uniforme de α', para qualquer $\varepsilon > 0$ dado, pode-se obter $\delta > 0$ tal que $|P| < \delta \Rightarrow |\alpha'(\eta_i) - \alpha'(\xi_i)| < \varepsilon/M(b-a)$. Então $|P| < \delta \Rightarrow \left|\sum_i f(\xi_i)[\alpha'(\eta_i) - \alpha'(\xi_i)](t_i - t_{i-1})\right| < \varepsilon$. Isto quer dizer que, quando $|P| \to 0$, o último somatório da igualdade (*) tem limite zero. Portanto

$$\int_a^b f(t)d\alpha = \lim_{|P|\to 0} \Sigma(P^*) = \lim_{|P|\to 0} \Sigma f(\xi_i)\alpha'(\xi_i)(t_i - t_{i-1}) = \\ = \int_a^b f(t)\alpha'(t)dt.$$

Uma majoração útil para a integral de Stieltjes é a seguinte:

Teorema 5. *Se $f\colon [a,b] \to \mathbb{R}$ é contínua e $\alpha\colon [a,b] \to \mathbb{R}$ é de variação limitada então*

$$\left|\int_a^b f(t)\,d\alpha\right| \leq M \cdot V$$

onde $M = \sup\{|f(t)|; a \leq t \leq b\}$ e V é a variação total de α.

Demonstração: Para toda partição pontilhada $P^* = (P, \xi)$ de $[a,b]$ temos

$$|\Sigma(P^*)| \leq \Sigma|f(\xi_i)|\,|\alpha(t_i) - \alpha(t_{i-1})| \leq M \cdot \Sigma|\alpha(t_i) - \alpha(t_{i-1})| \leq M \cdot V.$$

logo $\left|\int_a^b f(t)\,d\alpha\right| = \lim_{|P|\to 0} |\Sigma(P^*)| \leq M \cdot V$.

Concluimos nosso resumo com a proposição abaixo.

Mudança de Variável na Integral de Stieltjes. *Sejam $f\colon [c,d] \to \mathbb{R}$ contínua, $\alpha\colon [c,d] \to \mathbb{R}$ de variação limitada e $\varphi\colon [a,b] \to [c,d]$ um homeomorfismo. Então*

$$\int_c^d f(t)\,d\alpha = \int_a^b f\varphi(s)d(\alpha\varphi) \text{ se } \varphi \text{ é crescente, ou}$$

$$\int_c^d f(t)\,d\alpha = -\int_a^b f\varphi(s)d(\alpha\varphi) \text{ se } \varphi \text{ é decrescente.}$$

Observação: $f\varphi = f \circ \varphi$ e $\alpha\varphi = \alpha \circ \varphi$.

Demonstração: Seja φ crescente. Então $P = \{s_0 < s_1 < \cdots < s_k\} \subset [a,b]$ é uma partição se, e somente se, pondo $t_i = \varphi(s_i)$, $\varphi P = \{t_0 < \cdots < t_k\} \subset [c,d]$ é também uma partição. $P^* = (P,\xi)$ é uma partição pontilhada de $[a,b]$ se, e somente se, $\varphi P^* = (\varphi P, \varphi\xi)$, com $\varphi\xi = (\varphi(\xi_i))$, é uma partição pontilhada de $[c,d]$. É claro que $\Sigma(f,\alpha,\varphi P^*) = \Sigma(f\varphi,\alpha\varphi,P^*)$ e que $|\varphi P| \to 0$ se, e somente se, $|P| \to 0$. Daí resulta a primeira afirmação do teorema. No caso de φ ser decrescente, dada a partição $P = \{s_0 < s_1 < \cdots < s_k\}$ de $[a,b]$, pomos $t_i = \varphi(s_{k-i})$ e obtemos uma partição $\bar{\varphi}P = \{t_0 < \cdots < t_k\}$. Se os números ξ_i pontilham P, pontilharemos $\bar{\varphi}P$ com os números $\bar{\varphi}\xi_i = \varphi(\xi_{k-i})$. Vemos que $\Sigma(f,\alpha,\bar{\varphi}P^*) = -\Sigma(f\varphi,\alpha\varphi,P^*)$. Segue-se a segunda afirmação do teorema.

Observação: A fórmula análoga para integral de Riemann foi enunciada no Volume 1 como

$$\int_{\varphi(a)}^{\varphi(b)} f(t)\,dt = \int_a^b f\,\varphi(s)\cdot\varphi'(s)\,ds = \int_a^b f\,\varphi(s)\,d\varphi.$$

O sinal menos, no caso de φ ser decrescente, não aparece explicitamente aí mas acha-se oculto nos limites de integração, $\varphi(a)$ e $\varphi(b)$, da primeira integral. Sendo φ decrescente, tem-se $\varphi(a) = d$ e $\varphi(b) = c$, donde

$$\int_{\varphi(a)}^{\varphi(b)} f(t)\,dt = -\int_c^d f(t)\,dt.$$

3 Integral de uma forma ao longo de um caminho

Sejam $\omega = \sum_{i=1}^{n} a_i\,dx_i$ uma 1-forma definida num conjunto $X \subset \mathbb{R}^n$ e $\lambda\colon [a,b] \to X$ um caminho em $\mathbb{R}^n$, cujos valores $\lambda(t)$ pertencem ao conjunto X.

Definiremos a integral $\int_\lambda \omega$, da forma ω ao longo do caminho λ, como o limite

$$\int_\lambda \omega = \lim_{|P|\to 0} \sum_{j=1}^{k} \omega(\lambda(\xi_j))\cdot[\lambda(t_j) - \lambda(t_{j-1})],$$

onde $P^* = (P, \xi)$ é uma partição pontilhada de $[a, b]$, com $P = \{a = t_0 < t_1 < \cdots < t_k = b\}$ e $t_{j-1} \leq \xi_j \leq t_j$ para cada $j = 1, \ldots, k$.

Quando a forma ω (ou seja, cada função a_i) é contínua e o caminho λ é contínuo e retificável (isto é, cada função coordenada x_i é contínua e de variação limitada), o limite acima existe.

Com efeito, neste caso as integrais de Stieltjes $\int_a^b a_i(\lambda(t))\, dx_i$ existem, para $i = 1, \ldots, n$. Como

$$\omega(\lambda(\xi_j)) \cdot [\lambda(t_j) - \lambda(t_{j-1})] = \sum_{i=1}^{n} a_i(\lambda(\xi_j))[x_i(t_j) - x_i(t_{j-1})],$$

temos então

$$\int_\lambda \omega = \sum_{i=1}^{n} \int_a^b a_i(\lambda(t))\, dx_i\,.$$

Por analogia com a integral de Stieltjes na reta, escreve-se também $\int_\lambda \omega = \int_a^b \omega(\lambda(t)) \cdot d\lambda$.

Quando o caminho λ é de classe C^1, o Teorema 4 nos dá:

$$\int_\lambda \omega = \sum_{i=1}^{n} \int_a^b a_i(\lambda(t)) x_i'(t)\, dt = \int_a^b \omega(\lambda(t)) \cdot \lambda'(t)\, dt.$$

Isto reduz a integral curvilínea $\int_\lambda \omega$ à integral de Riemann da função $t \mapsto \omega(\lambda(t)) \cdot \lambda'(t)$, definida no intervalo $[a, b]$.

Para caminhos de classe C^1, as igualdades acima podem servir como definição de $\int_\lambda \omega$ mas, assim procedendo, perde-se o significado intuitivo da definição que demos.

Convém salientar, por outro lado, que a fórmula

$$\int_\lambda \omega = \int_a^b \omega(\lambda(t)) \cdot \lambda'(t)\, dt$$

é extremamente útil para os cálculos. Na apresentação que faremos da teoria das integrais curvilíneas, a classe de caminhos mais conveniente é a dos caminhos C^1 por partes, ou seccionalmente C^1. Esses caminhos são retificáveis. O caminho $\lambda\colon [a, b] \to \mathbb{R}^n$ é de classe C^1 por partes quando existe uma partição $P = \{a = t_0 < t_1 < \cdots < t_k = b\}$ do intervalo $[a, b]$ tal que cada restrição $\lambda_j = \lambda|[t_{j-1}, t_j],\ \ 1 \leq j \leq k$, é de classe C^1. Então

$$\int_\lambda \omega = \int_a^b \omega(\lambda(t)) \cdot d\lambda = \sum_j \int_{t_{j-1}}^{t_j} \omega(\lambda_j(t)) \cdot d\lambda_j = \sum_j \int_{\lambda_j} \omega$$

onde, para cada $j = 1, \ldots, k$, $\int_{\lambda_j} \omega = \int_{t_{j-1}}^{t_j} \omega(\lambda_j(t)) \cdot \lambda_j'(t)\, dt$.

A integral $\int_\lambda \omega$ é invariante por mudanças de parâmetro que mantenham o sentido de percurso do caminho λ. Por outro lado, uma reparametrização que inverta o sentido de λ troca o sinal de $\int_\lambda \omega$. Este é o conteúdo do

Teorema 6. *Sejam ω uma 1-forma contínua num conjunto $X \subset \mathbb{R}^n$, $\lambda\colon [c,d] \to X$ um caminho contínuo retificável em X e $\varphi\colon [a,b] \to [c,d]$ um homeomorfismo. Se φ for crescente (isto é, $\varphi(a) = c$, $\varphi(b) = d$) então $\int_\lambda \omega = \int_{\lambda\circ\varphi} \omega$. Se, porém, φ for decrescente, então $\int_{\lambda\circ\varphi} \omega = -\int_\lambda \omega$.*

Demonstração: Conseqüência imediata do teorema de mudança de variável na integral de Stieltjes.

Quando o caminho λ é de classe C^1 a demonstração abaixo, que usa apenas a mudança de variável clássica na integral de Riemann (Volume 1, pág. 326), vale para uma função $\varphi \in C^1$, mesmo que não seja monótona.

Teorema 6a. *Sejam ω uma 1-forma contínua num conjunto $X \subset \mathbb{R}^n$, $\lambda\colon [c,d] \to X$ um caminho de classe C^1 em X e $\varphi\colon [a,b] \to [c,d]$ uma função de classe C^1. Se $\varphi(a) = c$ e $\varphi(b) = d$ então $\int_\lambda \omega = \int_{\lambda\circ\varphi} \omega$. Se, porém, $\varphi(a) = d$ e $\varphi(b) = c$ então $\int_{\lambda\circ\varphi} \omega = -\int_\lambda \omega$.*

Demonstração: Quando $\varphi(a) = c$ e $\varphi(b) = d$, a fórmula clássica de mudança de variáveis na integral de Riemann dá

$$\int_\lambda \omega = \int_c^d \omega(\lambda(t)) \cdot \lambda'(t)\, dt = \int_a^b \omega(\lambda\varphi(s)) \cdot \lambda'(\varphi(s))\varphi'(s)\, ds =$$
$$= \int_a^b \omega(\lambda\varphi(s)) \cdot (\lambda\varphi)'(s)\, ds = \int_{\lambda\circ\varphi} \omega.$$

Analogamente, se $c = \varphi(b)$ e $d = \varphi(a)$, pela mesma fórmula:

$$\int_\lambda \omega = \int_c^d (\omega \circ \lambda) \cdot \lambda' dt = \int_b^a \omega \circ (\lambda \circ \varphi) \cdot (\lambda \circ \varphi)' ds =$$
$$= -\int_a^b \omega \circ (\lambda \circ \varphi) \cdot (\lambda \circ \varphi)' ds = -\int_{\lambda\circ\varphi} \omega.$$

Teorema 7. *Sejam ω uma 1-forma contínua no conjunto $X \subset \mathbb{R}^n$ e $\lambda\colon [a,b] \to X$ um caminho contínuo retificável em X, de comprimento*

L. *Se* $|\omega(\lambda(t))| \leq M$ *para todo* $t \in [a,b]$ *então*

$$\left|\int_\lambda \omega\right| \leq M \cdot L.$$

Demonstração: Para cada $t \in [a,b]$ e cada vetor $v \in \mathbb{R}^n$, temos $|\omega(\lambda(t)) \cdot v| \leq M \cdot |v|$. Portanto, dada qualquer partição pontilhada $P^* = (P, \xi)$ de $[a,b]$, com $P = \{t_0, \ldots, t_k\}$ e $t_{j-1} \leq \xi_j \leq t_j$, vale

$$\begin{aligned}|\Sigma(P^*)| &\leq \sum_j |\omega(\lambda(\xi_j)) \cdot [\lambda(t_j) - \lambda(t_{j-1})]| \leq \\ &\leq M \cdot \sum_j |\lambda(t_j) - \lambda(t_{j-1})| \leq M \cdot L.\end{aligned}$$

Observação: $|\omega(x)|$ é a norma do funcional linear $\omega(x) = \Sigma a_i(x) dx_i$, ou seja $|\omega(x)| = \sup\{|\omega(x) \cdot v|; v \in \mathbb{R}^n, |v| = 1\}$. Quando se toma em $\mathbb{R}^n$ a norma euclidiana, $|\omega(x)| = \sqrt{\Sigma a_i(x)^2}$. Se usarmos em $\mathbb{R}^n$ a norma do máximo, teremos $|\omega(x)| = \Sigma |a_i(x)|$. Quando a norma adotada em $\mathbb{R}^n$ é a da soma, tem-se $|\omega(x)| = \max\{|a_i(x)|; i = 1, \ldots, n\}$. (Veja a tabela no final do Capítulo I.)

Justificando o nome adotado, mostraremos no teorema abaixo que a integral da forma elemento de ângulo $\delta\theta$ ao longo de um caminho $\lambda\colon [a,b] \to \mathbb{R}^2 - \{0\}$, de classe C^1 por partes, é a "variação líquida"do ângulo que $\lambda(t)$ faz com o eixo das abcissas quando t varia de a até b. Por "variação líquida", entendemos a variação positiva menos a variação negativa. O significado preciso desta terminologia é fornecido pelo

Teorema 8. *Dado um caminho* $\lambda\colon [a,b] \to \mathbb{R}^2 - \{0\}$, *de classe* C^1 *por partes, sejam* $\alpha\colon [a,b] \to \mathbb{R}$ *uma função de classe* C^1 *por partes tal que* $\lambda(t) = |\lambda(t)| \cdot (\cos\alpha(t), \operatorname{sen}\alpha(t))$ *para todo* $t \in [a,b]$. (Veja o Corolário 2 do Teorema 11, Capítulo II.) *Então*

$$\int_\lambda \delta\theta = \alpha(b) - \alpha(a).$$

Demonstração: Suponhamos que $\lambda(t) = (x(t), y(t))$ seja de classe C^1 em todo o intervalo $[a,b]$. O mesmo ocorre com α. Temos

$$\int_\lambda \delta\theta = \int_a^b \frac{-yx' + xy'}{x^2 + y^2}\, dt.$$

Como $x = \rho\cos\alpha$ e $y = \rho\,\mathrm{sen}\,\alpha$, onde $\rho = \rho(t) = |\lambda(t)|$, um cálculo elementar fornece $(-yx' + xy')/(x^2+y^2) = \alpha'$, logo

$$\int_\lambda \delta\theta = \int_a^b \alpha'(t)\,dt = \alpha(b) - \alpha(a).$$

No caso geral, temos uma partição $P = \{a = t_0 < t_1 < \cdots < t_k = b\}$ tal que, para cada $j = 1,\ldots,k$, as restrições $\lambda_j = \lambda|[t_{j-1},t_j]$ e $\alpha|[t_{j-1},t_j]$ são de classe C^1. Então

$$\int_\lambda \delta\theta = \sum_{j=1}^k \int_{\lambda_j} \delta\theta = \sum_{j=1}^k [\alpha(t_j) - \alpha(t_{j-1})] = \alpha(b) - \alpha(a).$$

Corolário. *Se o caminho $\lambda\colon [a,b] \to \mathbb{R}^2 - \{0\}$, C^1 por partes, é fechado (isto é, $\lambda(a) = \lambda(b)$) então o número*

$$\frac{1}{2\pi}\int_\lambda \delta\theta = \frac{1}{2\pi}\int_\lambda \frac{-y\,dx + x\,dy}{x^2+y^2}$$

é um inteiro (positivo, negativo ou zero).

Com efeito, sendo $\lambda(a) = \lambda(b)$, temos $\cos\alpha(a) = \cos\alpha(b)$ e $\mathrm{sen}\,\alpha(a) = \mathrm{sen}\,\alpha(b)$, logo $\alpha(b) - \alpha(a)$ é um múltiplo inteiro de 2π, portanto $\dfrac{1}{2\pi}\displaystyle\int_\lambda \delta\theta = \dfrac{\alpha(b)-\alpha(a)}{2\pi}$ é um número inteiro.

Quando $\lambda\colon [a,b] \to \mathbb{R}^2 - \{0\}$ é um caminho fechado, de classe C^1 por partes, o número inteiro $\dfrac{1}{2\pi}\displaystyle\int_\lambda \delta\theta$ chama-se o *número de voltas* que o caminho λ descreve em torno da origem $0 \in \mathbb{R}^2$, quando t varia de a até b.

4 Justaposição de caminhos: caminho inverso

Dado um caminho $\lambda\colon [a,b] \to \mathbb{R}^n$, escrevemos $\lambda = \lambda_1 + \lambda_2$, e dizemos que λ é obtido por *justaposição* de λ_1 e λ_2 quando, para algum $c \in [a,b]$, tem-se $\lambda_1 = \lambda|[a,c]$ e $\lambda_2 = \lambda|[c,b]$.

De acordo com esta definição, para que tenha sentido a justaposição $\lambda_1 + \lambda_2$, é necessário (e suficiente) que os intervalos de definição de λ_1 e λ_2 tenham um ponto c em comum, sendo além disso $\lambda_1(c) = \lambda_2(c)$. A

ordem das "parcelas"em $\lambda_1 + \lambda_2$ é importante: o ponto c é o último no domínio de λ_1 e o primeiro no domínio de λ_2.

Analogamente, dado $\lambda\colon [a,b] \to \mathbb{R}^n$, se $P = \{t_0, \ldots, t_k\}$ é uma partição de $[a,b]$ e, para cada $j = 1, \ldots, k$, $\lambda_j = \lambda|[t_{j-1}, t_j]$, então $\lambda = \lambda_1 + \cdots + \lambda_k$, ou seja, λ é a justaposição dos caminhos $\lambda_1, \ldots, \lambda_k$, nesta ordem.

Se $\lambda = \lambda_1 + \lambda_2$, então λ é retificável se, e somente se, λ_1 e λ_2 são retificáveis. Analogamente, λ é de classe C^1 por partes se, e somente se, λ_1 e λ_2 também o são. Daí o interesse dos caminhos C^1 por partes: a justaposição de dois caminhos C^1 pode não ser C^1. Na realidade, λ é C^1 por partes se, e somente se, $\lambda = \lambda_1 + \cdots + \lambda_k$, onde cada λ_j é de classe C^1.

Quando $\lambda = \lambda_1 + \lambda_2$ é contínuo e retificável, com valores no domínio de uma forma contínua ω, então

$$\int_\lambda \omega = \int_a^b \omega(\lambda(t)) \cdot d\lambda = \int_a^c \omega(\lambda(t)) \cdot d\lambda + \int_c^b \omega(\lambda(t)) \cdot d\lambda =$$
$$= \int_{\lambda_1} \omega + \int_{\lambda_2} \omega.$$

Definiremos o *inverso* do caminho $\lambda\colon [a,b] \to \mathbb{R}^n$ como o caminho $-\lambda\colon [a,b] \to \mathbb{R}^n$, dado por $-\lambda(t) = \lambda(a+b-t)$. A função $\varphi\colon [a,b] \to [a,b]$, $\varphi(t) = a+b-t$, sendo um homeomorfismo decrescente, segue-se do Teorema 6 que, para toda forma ω, contínua num subconjunto de $\mathbb{R}^n$ que contenha a imagem do caminho contínuo e retificável λ, vale

$$\int_{-\lambda} \omega = -\int_\lambda \omega.$$

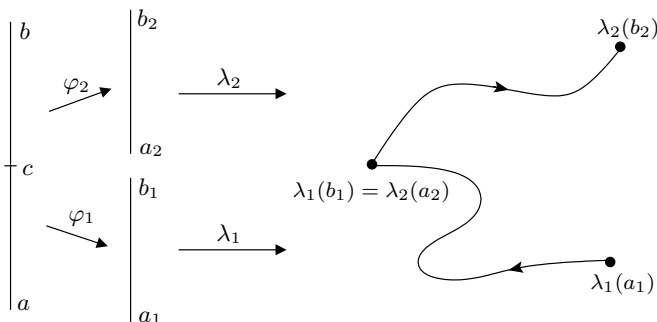
Evidentemente, λ contínuo e retificável, C^1 ou C^1 por partes implica que $-\lambda$ goza da mesma propriedade.

Vamos agora estender a definição de $\lambda_1 + \lambda_2$, permitindo justapor dois caminhos $\lambda_1\colon [a_1,b_1] \to \mathbb{R}^n$, $\lambda_2\colon [a_2,b_2] \to \mathbb{R}^n$ quando se tem apenas $\lambda_1(b_1) = \lambda_2(a_2)$, sem que seja necessariamente $b_1 = a_2$, como antes.

Diremos que os caminhos $\lambda\colon [c,d] \to \mathbb{R}^n$ e $\mu\colon [a,b] \to \mathbb{R}^n$ *diferem por uma reparametrização direta*, e escreveremos $\lambda \equiv \mu$, quando existir um homeomorfismo crescente $\varphi\colon [a,b] \to [c,d]$ tal que $\mu = \lambda \circ \varphi$.

A relação $\lambda \equiv \mu$ é reflexiva, simétrica e transitiva.

Se λ e μ são contínuos retificáveis, então $\lambda \equiv \mu \Rightarrow \int_\lambda \omega = \int_\mu \omega$ para toda forma contínua num conjunto que contenha a imagem de λ (igual à imagem de μ).

Quando $\mu = \lambda \circ \varphi$, onde $\varphi\colon [a,b] \to [c,d]$ é um homeomorfismo decrescente, então $-\lambda \equiv \mu$.

Sejam agora $\lambda_1\colon [a_1,b_1] \to \mathbb{R}^n$, $\lambda_2\colon [a_2,b_2] \to \mathbb{R}^n$ caminhos tais que $\lambda_1(b_1) = \lambda_2(a_2)$. Obteremos caminhos $\mu_1 \equiv \lambda_1$ e $\mu_2 \equiv \lambda_2$, para os quais faz sentido a justaposição $\lambda = \mu_1 + \mu_2$. Basta tomar um intervalo $[a,b]$, um ponto c, com $a < c < b$, homeomorfismos lineares crescentes $\varphi_1\colon [a,c] \to [a_1,b_1]$, $\varphi_2\colon [c,b] \to [a_2,b_2]$ e pôr $\mu_1 = \lambda_1 \circ \varphi_1$, $\mu_2 = \lambda_2 \circ \varphi_2$, $\lambda = \mu_1 + \mu_2$.

Convencionaremos escrever $\lambda = \lambda_1 + \lambda_2$ ainda neste caso, justificados pelo fato de valer a igualdade

$$\int_\lambda \omega = \int_{\lambda_1} \omega + \int_{\lambda_2} \omega$$

se ω é qualquer 1-forma contínua num conjunto que contenha as imagens dos caminhos contínuos retificáveis λ_1 e λ_2.

Novamente, se λ_1 e λ_2 são contínuos retificáveis, ou de classe C^1 por partes, o mesmo ocorre com $\lambda = \lambda_1 + \lambda_2$.

Em particular, como o ponto terminal do caminho λ coincide com o inicial de $-\lambda$, podemos formar o caminho justaposto $\lambda + (-\lambda)$ e obtemos

$$\int_{\lambda+(-\lambda)} \omega = \int_\lambda \omega + \int_{-\lambda} \omega = 0$$

para toda 1-forma ω, contínua num conjunto que contenha a imagem do caminho contínuo retificável λ.

Quando λ_1 e λ_2 são caminhos contínuos retificáveis *fechados* e tem sentido tomar $\lambda_1 + \lambda_2$ (o que significa que os quatro pontos extremos dos intervalos de definição de λ_1 e λ_2 são transformados por esses caminhos num único ponto de $\mathbb{R}^n$) então tem sentido também falar em $\lambda_2 + \lambda_1$. Se $\lambda_1 \neq \lambda_2$ então $\lambda_1 + \lambda_2 \neq \lambda_2 + \lambda_1$. Mas, para toda forma ω, contínua num conjunto que contenha as imagens de λ_1 e λ_2 , vale

$$\int_{\lambda_1+\lambda_2} \omega = \int_{\lambda_1} \omega + \int_{\lambda_2} \omega = \int_{\lambda_2} \omega + \int_{\lambda_1} \omega = \int_{\lambda_2+\lambda_1} \omega.$$

5 Integral curvilínea de um campo de vetores e de uma função

Além das formas diferenciais de grau 1, também os campos vetoriais podem ser integrados ao longo de caminhos, conforme descreveremos agora.

Um *campo de vetores* F num subconjunto $X \subset \mathbb{R}^n$ é simplesmente uma aplicação $F\colon X \to \mathbb{R}^n$. A terminologia se deve a exemplos físicos como campo gravitacional, campo eletrostático, campo de velocidades de um fluido que escoa, etc.

Sabemos que a cada vetor $F \in \mathbb{R}^n$ corresponde um funcional linear $\omega_F \in (\mathbb{R}^n)^*$ tal que $\omega_F \cdot v = \langle F, v\rangle$ para todo $v \in \mathbb{R}^n$. Daí resulta que a todo campo de vetores $F\colon X \to \mathbb{R}^n$ se pode fazer corresponder uma 1-forma $\omega_F\colon X \to (\mathbb{R}^n)^*$ tal que, para todo $x \in X$ e todo $v \in \mathbb{R}^n$, se tem $\omega_F(x) \cdot v = \langle F(x), v\rangle$.

Se, para cada $x \in X$, tem-se $F(x) = (a_1(x), \ldots, a_n(x))$ então $\omega_F = \sum\limits_{i=1}^{n} a_i\, dx_i$.

Dado um campo de vetores contínuo $F\colon X \to \mathbb{R}^n$, para todo caminho contínuo retificável $\lambda\colon [a, b] \to X$ se pode definir a integral de F ao longo de λ pondo:

$$\int_\lambda F = \int_\lambda \omega_F = \lim_{|P|\to 0} \sum_j \langle F(\lambda(\xi_j)), \lambda(t_j) - \lambda(t_{j-1})\rangle.$$

Quando F é um campo de forças, cada parcela do somatório acima é o "trabalho elementar"que uma força constante, igual a $F(\lambda(\xi_j))$ executaria ao longo do segmento de reta $[\lambda(t_{j-1}), \lambda(t_j)]$. A passagem ao limite justifica que se interprete a integral $\int_\lambda F$ como o *trabalho* realizado pela força F ao longo do caminho λ.

Se o caminho λ é de classe C^1 então

$$\int_\lambda F = \int_a^b \langle F(\lambda(t)), \lambda'(t)\rangle\, dt.$$

Evidentemente, a integral de um campo de vetores ao longo de um caminho goza de propriedades inteiramente análogas às da integral de uma 1-forma.

Sejam agora $f\colon X \to \mathbb{R}$ uma função contínua num conjunto $X \subset \mathbb{R}^n$ e $\lambda\colon [a,b] \to X$ um caminho contínuo retificável cuja imagem está contida em X. Definiremos a *integral de* f ao longo de λ pondo

$$\int_\lambda f\, ds = \lim_{|P|\to 0} \sum_j f(\lambda(\xi_j)) \cdot |\lambda(t_j) - \lambda(t_{j-1})|.$$

Quando λ é de classe C^1, temos

$$\int_\lambda f\, ds = \int_a^b f(\lambda(t)) \cdot |\lambda'(t)|\, dt.$$

Se $\varphi\colon [c,d] \to [a,b]$ for um difeomorfismo crescente de classe C^1, isto é, $\varphi'(t) > 0$ para todo $t \in [c,d]$, então $\int_\lambda f ds = \int_{\lambda\circ\varphi} f ds$. Do mesmo modo, quando for $\varphi'(t) < 0$ para todo $t \in [c,d]$ então $\varphi(c) = b$, $\varphi(d) = a$ e, como resulta da fórmula de mudança de variáveis, temos ainda $\int_\lambda f ds = \int_{\lambda\circ\varphi} f ds$.

Com este significado, diz-se que a integral $\int_\lambda f ds$ *independe* do sentido de percurso do caminho λ. (Ao contrário de $\int_\lambda \omega$ e $\int_\lambda F$, que são integrais "orientadas".)

Exemplos de grandezas físicas expressas por uma integral do tipo $\int_\lambda f ds$ são: a massa de um fio cuja densidade f é conhecida em cada um dos seus pontos, o momento de inércia do mesmo fio em torno de um eixo ou as coordenadas do vetor força gravitacional com que esse fio atrai um ponto de massa m_0.

6 Formas exatas e formas fechadas

Seja $U \subset \mathbb{R}^n$ um aberto. Uma forma diferencial $\omega\colon U \to (\mathbb{R}^n)^*$ diz-se *exata* em U quando existe uma função diferenciável $f\colon U \to \mathbb{R}$ tal que $\omega = df$.

Se $\omega = \Sigma a_i dx_i$, isto significa que $a_i = \partial f/\partial x_i$ para $i = 1, \ldots, n$.

A função $f\colon U \to \mathbb{R}$, tal que $df = \omega$, chama-se uma *primitiva* da forma ω. Se U é conexo, duas primitivas da mesma forma f diferem por uma constante.

A forma $\omega = \Sigma a_i dx_i \colon U \to (\mathbb{R}^n)^*$ é exata se, e somente se, o campo de vetores $F = (a_1, \ldots, a_n)$ é o gradiente de uma função $f\colon U \to \mathbb{R}$. Neste caso, a função f chama-se o *potencial* do campo F.

Note a importância do domínio U para o conceito de exatidão de uma forma. Para que ω seja chamada uma forma exata em U, a função f da qual ela é diferencial deve estar definida em todo o conjunto U e a igualdade $df(x) = \omega(x)$ deve verificar-se para cada ponto $x \in U$. Por exemplo, vimos no §1 e resulta imediatamente do teorema abaixo que a forma $\delta\theta$, elemento de ângulo, não é exata em $U = \mathbb{R}^2 - \{0\}$ mas, para toda semi-reta fechada ρ, partindo da origem, $\delta\theta$ é exata em $V_\rho = \mathbb{R}^2 - \rho$.

Teorema 9. *Seja $\omega\colon U \to (\mathbb{R}^n)^*$ uma 1-forma contínua no aberto $U \subset \mathbb{R}^n$. As seguintes afirmações são equivalentes:*

1. *ω é exata em U;*

2. *A integral $\int_\lambda \omega$ ao longo de um caminho $\lambda\colon [a,b] \to U$, seccionalmente C^1, depende apenas das extremidades $\lambda(a)$ e $\lambda(b)$;*

3. *Para todo caminho fechado $\lambda\colon [a,b] \to U$, seccionalmente C^1, tem-se $\int_\lambda \omega = 0$.*

Demonstração: Provemos que $1 \Leftrightarrow 2$. Se $\omega = df$ então

$$\int_\lambda \omega = \int_a^b df(\lambda(t)) \cdot \lambda'(t)dt = \int_a^b (f \circ \lambda)'(t)dt = f(\lambda(b)) - f(\lambda(a))$$

para todo caminho $\lambda\colon [a,b] \to U$, de classe C^1. Daí resulta, por aditividade, que $\int_\lambda \omega = f(\lambda(b)) - f(\lambda(a))$ se λ é seccionalmente C^1, logo $\int_\lambda \omega$ depende apenas das extremidades $\lambda(a)$, $\lambda(b)$ do caminho λ. Em seguida, provamos que $2 \Rightarrow 1$. Suponhamos inicialmente que U seja conexo. Fixemos um ponto $p \in U$. Para cada $x \in U$, existe um caminho seccionalmente C^1, $\lambda\colon [a,b] \to U$, com $\lambda(a) = p$ e $\lambda(b) = x$. A integral $\int_\lambda \omega$, por hipótese, depende apenas de p e x, logo podemos indicá-la com $\int_p^x \omega$ e definir uma função $f\colon U \to \mathbb{R}$ pondo $f(x) = \int_p^x \omega$. Afirmamos que $df = \omega$, isto é, se $\omega = \Sigma a_i dx_i$ então $\dfrac{\partial f}{\partial x_i} = a_i$ para todo $i = 1, \ldots, n$.

Com efeito, tomando sempre a derivada $\frac{d}{dt}$ no ponto $t = 0$, temos:

$$\begin{aligned}\frac{\partial f}{\partial x_i}(x) &= \frac{d}{dt}\, f(x+te_i) = \frac{d}{dt}\,[f(x+te_i) - f(x)] = \\ &= \frac{d}{dt}\int_x^{x+te_i} \omega = \frac{d}{dt}\int_0^t \omega(x+se_i)\cdot e_i\, ds,\end{aligned}$$

onde a última igualdade se deve a termos ligado os pontos x e $x+te_i$ pelo caminho $\lambda\colon [0,t] \to U$, $\lambda(s) = x + se_i$, com $\lambda'(s) = e_i$ para todo s. Como $\omega(y)\cdot e_i = a_i(y)$, continuando a tomar a derivada $\frac{d}{dt}$ no ponto $t = 0$, temos:

$$\frac{\partial f}{\partial x_i}(x) = \frac{d}{dt}\int_0^t a_i(x+se_i)ds = a_i(x).$$

No caso geral, basta observar que, para definir uma função diferenciável $f\colon U \to \mathbb{R}$, é suficiente definí-la (independentemente) em cada componente conexa de U.

A equivalência entre as afirmações 2 e 3 é imediata, pois se λ e μ têm as mesmas extremidades então $\lambda + (-\mu) = \lambda - \mu$ é um caminho fechado e, reciprocamente, todo caminho fechado v se escreve (de infinitas maneiras) sob a forma $v = \lambda - \mu$, onde λ e μ são caminhos com as mesmas extremidades. Assim, $\int_\lambda \omega = \int_\mu \omega \Leftrightarrow \int_{\lambda-\mu}\omega = 0 \Leftrightarrow \int_v \omega = 0$.

Como aplicação do teorema acima, vemos, mais uma vez, que a forma elemento de ângulo, $\delta\theta$, não é exata em $U = \mathbb{R}^2 - \{0\}$ pois $\lambda\colon [0,2\pi] \to U$, definido por $\lambda(t) = (\cos t, \operatorname{sen} t)$, é um caminho fechado em U, tal que $\int_\lambda \delta\theta = 2\pi \neq 0$.

Abordaremos agora o problema de obter condições mediante as quais uma forma $\omega = \Sigma a_i dx_i\colon U \to (\mathbb{R}^n)^*$ seja exata. Isto equivale a indagar quando o sistema de equações diferenciais parciais $\frac{\partial f}{\partial x_i} = a_i$ $(i = 1,\ldots,n)$ possui uma solução $f\colon U \to \mathbb{R}$. Uma condição necessária para a integrabilidade desse sistema resulta imediatamente do Teorema de Schwarz: se $\omega = \Sigma a_i dx_i$ é uma forma exata de classe C^1 então $\frac{\partial a_i}{\partial x_j} = \frac{\partial a_j}{\partial x_i}$ para quaisquer $i, j = 1,\ldots,n$. Com efeito, $\omega = df$ nos dá

$$\frac{\partial a_i}{\partial x_j} = \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right) = \frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right) = \frac{\partial a_j}{\partial x_i}\,,$$

pois f é, neste caso, de classe C^2.

Essa condição necessária não é suficiente para que a forma $\omega\colon U \to (\mathbb{R}^n)^*$ seja exata. Com efeito, escrevendo o elemento de ângulo como $\delta\theta = adx + bdy$, com $a = \dfrac{-y}{x^2+y^2}$ e $b = \dfrac{x}{x^2+y^2}$, temos $\dfrac{\partial a}{\partial y} = \dfrac{\partial b}{\partial x}$ mas $\delta\theta$ não é exata em $U = \mathbb{R}^2 - \{0\}$.

Diremos que uma forma $\omega = \Sigma a_i dx_i$, de classe C^1, é *fechada* quando $\dfrac{\partial a_i}{\partial x_j} = \dfrac{\partial a_j}{\partial x_i}\cdot$

Toda forma exata de classe C^1 é fechada. A recíproca é falsa (como se vê pelo elemento de ângulo) e, de fato, não seria razoável esperar que ela valesse, pois a propriedade de ser fechada, expressa em termos de derivadas, é apenas local.

Convencionemos dizer que uma forma $\omega\colon U \to (\mathbb{R}^n)^*$ é *localmente exata* quando todo ponto de U estiver contido num aberto, restrito ao qual ω é exata. Então vale o melhor possível, que é o

Teorema 10. *Uma forma $\omega\colon U \to (\mathbb{R}^n)^*$, de classe C^1, é fechada se, e somente se, é localmente exata.*

Demonstração: Toda forma C^1 localmente exata é obviamente fechada. Basta então provar que se ω é fechada numa bola B então ω é exata em B. Para simplificar a notação, suporemos que $\omega = adx + bdy + cdz$ é fechada na bola B, de centro na origem em $\mathbb{R}^3$. O caso geral se trata do mesmo modo. Nossa hipótese é que $\partial a/\partial y = \partial b/\partial x$, $\partial a/\partial z = \partial c/\partial x$ e $\partial b/\partial z = \partial c/\partial y$. Definimos $f\colon B \to \mathbb{R}$ pondo, para cada $(x, y, z) \in B$:

$$\begin{aligned} f(x,y,z) &= \int_0^1 [a(tx,ty,tz)x + b(tx,ty,tz)y + c(tx,ty,tz)z]dt = \\ &= \int_\lambda \omega, \end{aligned}$$

onde $\lambda\colon [0,1] \to B$ é dado por $\lambda(t) = (tx, ty, tz)$. Pela Regra de Leibniz (derivação sob o sinal de integral) temos

$$\frac{\partial f}{\partial x}(x,y,z) = \int_0^1 \left[\frac{\partial a}{\partial x}\cdot tx + a + \frac{\partial b}{\partial x}\cdot ty + \frac{\partial c}{\partial x}\cdot tz\right] dt,$$

onde acima (e no que se segue) as derivadas parciais de a, b e c são tomadas no ponto (tx, ty, tz). Levando em conta que $\partial b/\partial x = \partial a/\partial y$ e

$\partial c/\partial x = \partial a/\partial z$, vem

$$\frac{\partial f}{\partial x}(x,y,z) = \int_0^1 \left[\left(\frac{\partial a}{\partial x}\cdot x + \frac{\partial a}{\partial y}\cdot y + \frac{\partial a}{\partial z}\cdot z\right)t + a\right] dt =$$
$$= \int_0^1 [(a\circ\lambda)'t + a]dt = \int_0^1 [(a\circ\lambda)\cdot t]'\, dt = a(x,y,z).$$

De modo análogo se vê que $\partial f/\partial y = b(x,y,z)$ e $\partial f/\partial z = c(x,y,z)$, logo $df = \omega$.

Note que a hipótese de B ser uma bola intervém na demonstração acima para assegurar, quando $0 = (0,0,0) \in B$ e $p = (x,y,z) \in B$, que o segmento de reta $[0,p]$ está contido em B. Isto seria óbvio se B fosse qualquer conjunto convexo. Assim toda forma fechada cujo domínio é um aberto convexo no espaço $\mathbb{R}^n$ é uma forma exata.

Exemplo 1. Vejamos um problema que se reduz a determinar se uma certa forma fechada é ou não exata. Seja $u\colon U \to \mathbb{R}$ uma função *harmônica* num aberto $U \subset \mathbb{R}^2$, isto é, uma função de classe C^2 tal que $\partial^2 u/\partial x^2 + \partial^2 u/\partial y^2 = 0$ identicamente em U. (Exemplos: $u(x,y) = x^2 - y^2$ ou $u(x,y) = xy$, com $U = \mathbb{R}^2$, e $u(x,y) = \frac{1}{2}\log(x^2+y^2)$, com $U = \mathbb{R}^2 - \{0\}$.) Associemos à função u a forma $\omega\colon U \to (\mathbb{R}^2)^*$, dada por $\omega = -\frac{\partial u}{\partial y}dx + \frac{\partial u}{\partial x}dy$. A harmonicidade de U assegura que a forma ω é fechada. Ela será exata se, e somente se existir uma função $v\colon U \to \mathbb{R}$, de classe C^1 (e portanto C^2) tal que $\partial v/\partial x = -\partial u/\partial y$ e $\partial v/\partial y = \partial u/\partial x$. Ora, estas são as equações de Cauchy-Riemann, condições necessárias e suficientes para que seja holomorfa a função complexa $f\colon U \to \mathbb{C}$, dada por $f(x+iy) = u(x,y) + iv(x,y)$. (Veja Exemplo 6, Capítulo III.) Assim a forma fechada $\omega\colon U \to (\mathbb{R}^2)^*$ é exata se, e somente se, a função harmônica u é a parte real de uma função holomorfa $f = u + iv$, definida no domínio de u. Como toda forma fechada é localmente exata, segue-se que toda função harmônica num aberto do plano é localmente a parte real de uma função holomorfa. Globalmente, isto não é necessariamente verdade, como veremos agora. No caso da função $u\colon \mathbb{R}^2 \to \mathbb{R}$, $u(x,y) = x^2 - y^2$, obtemos a forma $\omega = 2ydx + 2xdy$, que é exata: temos $\omega = dv$, onde $v(x,y) = 2xy$. De fato, u é a parte real da função holomorfa $f\colon \mathbb{C} \to \mathbb{C}$, $f(z) = z^2$. Por outro lado, tomando a função harmônica $u\colon \mathbb{R}^2 - \{0\} \to \mathbb{R}$, $u(x,y) = \frac{1}{2}\log(x^2+y^2)$, obtemos a forma $\omega = (x^2+y^2)^{-1}(-ydx + xdy)$, ou seja, $\omega = \delta\theta$, logo ω não é exata em

$\mathbb{R}^2 - \{0\}$. Assim, a função harmônica $u = \dfrac{1}{2}\log(x^2+y^2)$ não é a parte real de nenhuma função holomorfa em $\mathbb{R}^2 - \{0\}$. Mas se ρ é qualquer semi-reta fechada partindo da origem, sabemos que ω é exata em $\mathbb{R}^2 - \rho$, logo existe $v\colon \mathbb{R}^2 - \rho \to \mathbb{R}$ tal que $dv = \omega$; assim $f = u + iv$ é holomorfa. [Em virtude do Teorema 2, subtraindo de v uma constante, se necessário, (o que não prejudicará a propriedade $dv = \omega$) podemos supor que v seja uma função-ângulo em $\mathbb{R}^2 - \rho$. Então, para todo $z = (x, y) \notin \rho$, teremos

$$\begin{aligned} e^{f(z)} &= e^{u+iv} = e^u \cdot e^{iv} = \sqrt{x^2+y^2}(\cos v + i \operatorname{sen} v) = \\ &= \sqrt{x^2+y^2}\left(\frac{x}{\sqrt{x^2+y^2}} + i\,\frac{y}{\sqrt{x^2+y^2}}\right) = z. \end{aligned}$$

A identidade $e^{f(z)} = z$ significa que $f\colon \mathbb{R}^2 - \rho \to \mathbb{C}$ é um "ramo do logaritmo".]

Exemplo 2. Ainda no contexto de variáveis complexas, seja $f = a + ib$ uma função holomorfa no aberto $U \subset \mathbb{C}$. Por Cauchy-Riemann, as duas formas diferenciais $\omega = adx - bdy$ e $\varphi = bdx + ady$ são fechadas em U. Estas formas são exatas em U se, e somente se, existem funções $u, v\colon U \to \mathbb{R}$ (necessariamente de classe C^2) tais que $\partial u/\partial x = a$, $\partial u/\partial y = -b$, $\partial v/\partial x = b$ e $\partial v/\partial y = a$. Então a função complexa $g = u + iv\colon U \to \mathbb{C}$ cumpre as equações de Cauchy-Riemann e tem-se $g' = \partial u/\partial x + i\cdot\partial v/\partial x = a + ib = f$, isto é, g é uma função holomorfa cuja derivada é f. Assim, para que a função holomorfa $f = a + ib\colon U \to \mathbb{C}$ possua uma primitiva definida em U (ou seja, uma função $g\colon U \to \mathbb{C}$ tal que $g' = f$) é necessário e suficiente que as duas formas diferenciais $\omega = adx - bdy$ e $\varphi = bdx + ady$ sejam ambas exatas em U. Por exemplo, se tomarmos $f\colon \mathbb{C} - \{0\} \to \mathbb{C}$, $f(z) = 1/z$, teremos $f(z) = x/(x^2+y^2) - iy/(x^2+y^2)$, logo $\omega = (x^2+y^2)^{-1}(xdx + ydy)$ e $\varphi = \delta\theta$. A forma ω é exata em $U = C - \{0\}$; de fato $\omega = du$, onde $u = \dfrac{1}{2}\log(x^2+y^2)$. Mas, como sabemos, $\varphi = \delta\theta$ não é exata em U, logo f não possui primitiva em U.

Os exemplos acima dão uma idéia da relevância das formas diferenciais de grau 1 em Análise Complexa.

7 Homotopia

Temos em mente analisar mais a fundo a questão de saber se uma forma $\omega\colon U \to (\mathbb{R}^n)^*$ é exata. Já vimos que uma condição necessária, quando ω é de classe C^1, é que seja fechada. Por isso nos restringiremos a formas de classe C^1 e indagaremos agora em que condições uma forma fechada ω é exata.

A resposta vai depender não somente de ω mas da natureza topológica do seu domínio U. Já temos uma indicação disso: quando $U \subset \mathbb{R}^n$ é um subconjunto aberto e convexo do espaço $\mathbb{R}^n$, então qualquer forma fechada $\omega\colon U \to (\mathbb{R}^n)^*$ é exata.

As propriedades topológicas de U que influem nessa questão são as que dizem respeito à possibilidade de deformar um caminho em U continuamente. Elas são tratadas na teoria da homotopia, da qual apresentaremos algumas noções agora.

Sejam $\lambda, \mu\colon [a,b] \to X$ caminhos no conjunto $X \subset \mathbb{R}^n$, com o mesmo domínio $[a,b]$ e com as mesmas extremidades, isto é, $\lambda(a) = \mu(a)$ e $\lambda(b) = \mu(b)$.

Uma *homotopia* entre λ e μ é uma aplicação contínua $H\colon$ $[a,b] \times I \to X$, onde $I = [0,1]$, tal que $H(s,0) = \lambda(s)$, $H(s,1) = \mu(s)$, $H(a,t) = \lambda(a) = \mu(a)$ e $H(b,t) = \lambda(b) = \mu(b)$ para quaisquer $s \in [a,b]$, $t \in [0,1]$.

Quando existe uma homotopia H entre λ e μ, dizemos que esses caminhos são *homotópicos* e escrevemos $H\colon \lambda \simeq \mu$ ou, mais simplesmente $\lambda \simeq \mu$.

Intuitivamente, uma homotopia $H\colon \lambda \simeq \mu$ é uma deformação contínua do caminho λ, dentro de X, que acontece numa unidade de tempo, começando com λ (para $t = 0$) e terminando com μ (quando $t = 1$). Durante essa deformação, as extremidades $\lambda(a) = \mu(a)$ e $\lambda(b) = \mu(b)$ permanecem fixas. As vezes é conveniente imaginar a homotopia como uma "família a um parâmetro"de caminhos $H_t\colon [a,b] \to X$, onde $H_t(s) = H(s,t)$, todos esses caminhos começando no ponto $\lambda(a) = \mu(a)$ e terminando no ponto $\lambda(b) = \mu(b)$. Os caminhos H_t dependem continuamente do parâmetro $t \in [0,1]$, no sentido de que $H\colon [a,b] \times [0,1] \to X$ é contínua.

Quando o caminho λ é fechado, isto é, $\lambda(a) = \lambda(b)$, chamaremos $\lambda(a)$ *o ponto básico de* λ. Todo caminho homotópico a λ é também fechado e toda homotopia mantém fixo o ponto básico, isto é, $H_t(a) = H_t(b) =$

$\lambda(a) = \lambda(b)$, para todo $t \in [0,1]$.

Convém ressaltar que, ao escrevermos $H\colon \lambda \simeq \mu$, fica subentendido um dado absolutamente essencial que é o conjunto X, dentro do qual a homotopia $H\colon [a,b] \times I \to X$ tem lugar. Aumentando-se o conjunto X, dois caminhos que não eram homotópicos podem passar a ser, como veremos a seguir. A relação de homotopia é uma equivalência, ou seja, ela é:

1. *Reflexiva*: $\lambda \simeq \lambda$;
2. *Simétrica*: Se $\lambda \simeq \mu$ então $\mu \simeq \lambda$;
3. *Transitiva*; Se $\lambda \simeq \mu$ e $\mu \simeq v$ então $\lambda \simeq v$.

De fato, $H\colon [a,b] \times I \to X$, definida por $H(s,t) = \lambda(s)$, é uma homotopia entre λ e λ. Além disso, se $H\colon \lambda \simeq \mu$ então $K\colon [a,b] \times I \to X$, definida por $K(s,t) = H(s, 1-t)$, é uma homotopia entre μ e λ. Finalmente, dadas $H\colon \lambda \simeq \mu$ e $K\colon \mu \simeq v$, definimos $L\colon [a,b] \times I \to X$ pondo $L(s,t) = H(s, 2t)$ se $0 \leq t \leq 1/2$ e $L(s,t) = K(s, 2t-1)$ se $1/2 \leq t \leq 1$. Então L é contínua e é uma homotopia entre λ e v.

Uma maneira bastante conveniente de construir homotopias decorre do

Lema 1. *Sejam $\lambda, \mu\colon [a,b] \to X$ caminhos no conjunto $X \subset \mathbb{R}^n$. Se λ e μ possuem as mesmas extremidades e, além disso, para cada $s \in [a,b]$, o segmento de reta $[\lambda(s), \mu(s)]$ está contido em X então λ e μ são homotópicos.*

Definiremos, neste caso, uma homotopia $H\colon \lambda \simeq \mu$ pondo $H(s,t) = (1-t)\lambda(s) + t\mu(s)$, com $a \leq s \leq b$, $0 \leq t \leq 1$. Note que a homotopia H deforma λ em μ fazendo, para cada $s \in [a,b]$, o ponto $\lambda(s)$ deslizar, no tempo unitário, com velocidade constante, ao longo do segmento de reta $[\lambda(s), \mu(s)]$. Uma tal homotopia diz-se *linear* e os caminhos λ, μ dizem-se, então, *linearmente homotópicos.*

Por exemplo, se $X \subset \mathbb{R}^n$ é convexo, então dois caminhos quaisquer $\lambda, \mu\colon [a,b] \to X$, com as mesmas extremidades, são linearmente homotópicos.

Demonstraremos a seguir dois lemas que simplificarão bastante a prova do Teorema 11, um dos resultados principais deste capítulo.

Um caminho λ diz-se *poligonal* quando se obtém por justaposição $\lambda = \lambda_1 + \cdots + \lambda_k$ de caminhos retilíneos.

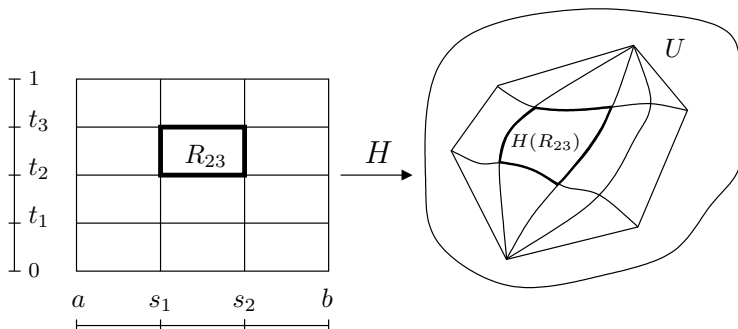
Lema 2. *Seja $U \subset \mathbb{R}^n$ aberto. Todo caminho $\lambda\colon [a,b] \to U$ é linearmente homotópico a um caminho poligonal.*

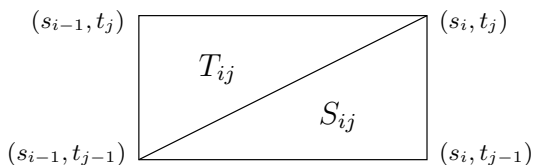
Demonstração: Pelo Teorema 29, Capítulo I, existe uma partição $a = s_0 < s_1 < \cdots < s_k = b$ tal que, para cada $i = 1, \ldots, k$, a imagem $\lambda([s_{i-1}, s_i])$ está contida numa bola aberta $B_i \subset U$, a qual contém também (por ser convexa) a imagem do caminho retilíneo $\mu_i \colon [s_{i-1}, s_i] \to U$, cujos extremos são $\lambda(s_{i-1})$ e $\lambda(s_i)$. Como cada bola B_i é convexa, segue-se do Lema 1 que cada restrição $\lambda|[s_{i-1}, s_i]$ é linearmente homotópica a μ_i, logo λ é linearmente homotópico ao caminho poligonal $\mu = \mu_1 + \cdots + \mu_k \colon [a, b] \to U$.

Continuaremos escrevendo $I = [0, 1]$.

Seja $H \colon [a, b] \times I \to X \subset \mathbb{R}^n$ uma homotopia. Para cada $s \in [a, b]$ e cada $t \in I$, indicaremos com $H^s \colon I \to X$ e $H_t \colon [a, b] \to X$, respectivamente, os caminhos definidos por $H^s(t) = H(s, t)$ e $H_t(s) = H(s, t)$. Diremos que a homotopia H é *poligonal* quando, para cada $s \in [a, b]$ e cada $t \in I$, o caminho "vertical" H^s e o caminho "horizontal" H_t forem poligonais.

Lema 3. *Seja $U \subset \mathbb{R}^n$ um aberto. Se dois caminhos poligonais $\lambda, \mu \colon [a, b] \to U$ são homotópicos, então existe uma homotopia poligonal entre λ e μ.*

Demonstração: Pelo Teorema 29, Capítulo I, existem partições $P = \{a = s_0 < s_1 < \cdots < s_k = b\}$ e $Q = \{0 = t_0 < t_1 < \cdots < t_r = 1\}$, as quais determinam uma decomposição do retângulo $[a, b] \times I$ em kr retângulos justapostos $R_{ij} = [s_{i-1}, s_i] \times [t_{j-1}, t_j]$, tais que a imagem $H(R_{ij})$ de cada um deles está contida numa bola $B_{ij} \subset U$. Então, $x, y \in H(R_{ij}) \Rightarrow [x, y] \subset U$. Refinando a partição P, se necessário, podemos supor que, em cada um dos

segmentos $[s_{i-1}, s_i] \times 0$ e $[s_{i-1}, s_i] \times 1$, $(i = 1, 2, \ldots, k)$, H é retilíneo. Definiremos a homotopia poligonal $K \colon [a, b] \times I \to U$, entre λ e μ, pri-

meiro nos vértices (s_i, t_j) da decomposição retangular de $[a, b] \times I$, pondo $K(s_i, t_j) = H(s_i, t_j)$. Em seguida, definiremos K em cada retângulo R_{ij} da decomposição assim: por meio da diagonal que liga os vértices (s_{i-1}, t_{j-1}) e (s_i, t_j) do retângulo R_{ij},

dividimos este retângulo em 2 triângulos, chamando de S_{ij} o inferior e de T_{ij} o superior. A aplicação K, já definida nos vértices de R_{ij} , será estendida a todo esse retângulo exigindo-se que ela seja afim em T_{ij} e em S_{ij} separadamente. Como uma aplicação afim transforma segmentos de reta em segmentos de reta, a homotopia K será poligonal.

Observação: Seja $T \subset \mathbb{R}^2$ um triângulo de vértices a, b, c. É fácil ver que T é o conjunto dos pontos $p \in \mathbb{R}^2$ que se escrevem (de modo único) sob a forma $p = ua + vb + wc$ onde $u, v, w \in [0, 1]$ e $u + v + w = 1$. [Os números u, v, w são chamados as "coordenadas baricêntricas"de p em T.] Uma aplicação $f : T \to \mathbb{R}^n$ diz-se *afim* quando, para $p, q \in T$ e $0 \leq t \leq 1$ quaisquer, tem-se $f((1-t)p+tq) = (1-t)f(p)+tf(q)$. Dados arbitrariamente os pontos $A, B, C \in \mathbb{R}^n$, existe uma única aplicação afim $f : T \to \mathbb{R}^n$ tal que $f(a) = A$, $f(b) = B$, $f(c) = C$. Basta pôr $f(ua + vb + wc) = uA + vB + wC$. (Veja os detalhes no apêndice a este capítulo.)

Corolário. *Se $\lambda, \mu : [a, b] \to U$ são caminhos de classe C^1 por partes, homotópicos no aberto $U \subset \mathbb{R}^n$, então existe uma homotopia $H : \lambda \simeq \mu$ tal que, para cada $s \in [a, b]$ e cada $t \in I$, os caminhos $H^s : I \to U$ e $H_t : [a, b] \to U$, definidos por $H^s(t) = H(s, t)$ e $H_t(s) = H(s, t)$ são de classe C^1 por partes.*

Com efeito, existem homotopias lineares $\lambda \simeq \lambda_1$ e $\mu \simeq \mu_1$, onde λ_1 e μ_1 são caminhos poligonais em U. Em seguida, existe uma homotopia poligonal $\lambda_1 \simeq \mu_1$. Compondo essas homotopias obtemos, por transitividade, uma homotopia $H : \lambda = \mu$, com a propriedade requerida.

8 Integrais curvilíneas e homotopia

Teorema 11. *Seja ω uma 1-forma fechada num aberto $U \subset \mathbb{R}^n$. Se $\lambda, \mu\colon [a,b] \to U$ são caminhos seccionalmente C^1, homotópicos em U, então $\int_\lambda \omega = \int_\mu \omega$.*

Demonstração: Seja $H\colon [a,b] \times I \to U$ uma homotopia entre λ e μ. Pelo Teorema 29, Capítulo I, existe uma decomposição do retângulo $[a,b] \times I$ em kr retângulos justapostos $R_{ij} = [s_{i-1}, s_i] \times [t_{j-1}, t_j]$, com $a = s_0 < s_1 < \cdots < s_k = b$ e $0 = t_0 < t_1 < \cdots < t_r = 1$, tais que cada imagem $H(R_{ij})$ está contida numa bola $B_{ij} \subset U$. Isto define os caminhos "horizontais"$\alpha_0, \alpha_1, \dots, \alpha_r\colon [a,b] \to U$, onde $\alpha_j(s) = H(s, t_j)$, logo $\alpha_0 = \lambda$ e $\alpha_r = \mu$; e os caminhos "verticais"$\beta_0, \beta_1, \dots, \beta_k\colon I \to U$, dados por $\beta_i(t) = H(s_i, t)$, resultando que β_0 e β_k são constantes. Pelo corolário no fim do parágrafo anterior, podemos supor que os caminhos α_j e β_i são todos seccionalmente C^1. Devemos provar, sucessivamente, as igualdades

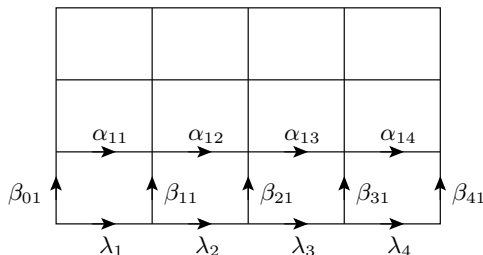
$$\int_\lambda \omega = \int_{\alpha_1} \omega, \int_{\alpha_1} \omega = \int_{\alpha_2} \omega, \dots, \int_{\alpha_{r-1}} \omega = \int_\mu \omega.$$

Basta demonstrar a primeira delas; as demais são análogas. Temos

$$\lambda = \lambda_1 + \cdots + \lambda_k\,, \text{ onde } \lambda_i = \lambda|[s_{i-1}, s_i]$$

e

$$\alpha_1 = \alpha_{11} + \cdots + \alpha_{1k}\,, \text{ onde } \alpha_{1i} = \alpha_1|s_{i-1}, s_i].$$

Escrevamos ainda $\beta_{i1} = \beta_i|[0, t_1]$, $i = 0, 1, \dots, k$. Assim, β_{01} e β_{k1} são caminhos constantes.

A forma fechada ω é exata em cada bola B_{ij}, logo suas integrais ao longo de dois caminhos com as mesmas extremidades, contidos em B_{ij},

são iguais. Assim,

$$\int_{\lambda_1} \omega = \int_{\beta_{01}+\alpha_{11}-\beta_{11}} \omega = \int_{\beta_{01}} \omega + \int_{\alpha_{11}} \omega - \int_{\beta_{11}} \omega = \\ = \int_{\alpha_{11}} \omega - \int_{\beta_{11}} \omega,$$

pois λ_1 e $\beta_{01} + \alpha_{11} - \beta_{11}$ são caminhos com as mesmas extremidades, contidos em B_{11}. Analogamente,

$$\int_{\lambda_2} \omega = \int_{\beta_{11}} \omega + \int_{\alpha_{12}} \omega - \int_{\beta_{21}} \omega,$$

....

....

$$\int_{\lambda_k} \omega = \int_{\beta_{k-1,1}} \omega + \int_{\alpha_{1k}} \omega - \int_{\beta_{k1}} \omega = \int_{\beta_{k-1,1}} \omega + \int_{\alpha_{1k}} \omega.$$

Daí, $\displaystyle\int_{\lambda} \omega = \sum_{i=1}^{k} \int_{\lambda_i} \omega = \sum_{i=1}^{k} \int_{\alpha_{1i}} \omega = \int_{\alpha_1} \omega$. Isto conclui a demonstração.

Mostraremos agora que, quando se trata de caminhos *fechados* $\lambda, \mu\colon [a,b] \to U \subset \mathbb{R}^n$, a validez do teorema acima não necessita que o ponto básico $\lambda(a) = \lambda(b) = \mu(a) = \mu(b)$ permaneça fixo durante a deformação. Isto se faz mediante a noção de homotopia livre entre caminhos fechados.

Uma *homotopia livre* entre os caminhos fechados $\lambda, \mu\colon [a,b] \to X \subset \mathbb{R}^n$ é uma aplicação contínua $H\colon [a,b] \times I \to X$ tal que $H(s,0) = \lambda(s)$, $H(s,1) = \mu(s)$ e $H(a,t) = H(b,t)$ para quaisquer $s \in [a,b]$ e $t \in I$. A última igualdade significa que, para todo $t \in I$, o caminho $H_t\colon [a,b] \to X$, definido por $H_t(s) = H(s,t)$, é fechado.

Diremos que os caminhos fechados $\lambda, \mu\colon [a,b] \to X$ são *livremente homotópicos* quando existir uma homotopia livre H entre λ e μ. Escreveremos então $H\colon \lambda \approx \mu$ ou, simplesmente, $\lambda \approx \mu$. A relação $\lambda \approx \mu$ é uma equivalência entre os caminhos fechados em X.

O teorema a que nos referimos é o seguinte:

Teorema 12. *Seja ω uma forma fechada num aberto $U \subset \mathbb{R}^n$. Se $\lambda, \mu\colon [a,b] \to U$ são caminhos fechados, seccionalmente C^1, livremente homotópicos em U, então $\int_\lambda \omega = \int_\mu \omega$.*

A demonstração é quase igual à do caso anterior, diferindo dela apenas porque, como H é uma homotopia livre, os caminhos "verticais" β_0

e β_k não são mais constantes. Em vez disso, como H_t é um caminho fechado para todo t, temos $\beta_0 = \beta_k$ e portanto $\beta_{01} = \beta_{k1}$. Isto acarreta apenas uma mudança: no cálculo final, as parcelas $\int_{\beta_{01}} \omega$ e $-\int_{\beta_{k1}} \omega$, em vez de serem ambas iguais a zero como no caso anterior, se cancelam na soma.

A noção de homotopia livre dá, *para caminhos fechados*, uma generalidade bem maior à igualdade $\int_\lambda \omega = \int_\mu \omega$ para ω fechada e λ, μ homotópicas, pois agora os caminhos λ e μ nem sequer precisam ter pontos em comum.

Por exemplo, dados $r_1 > 0$ e $r_2 > 0$ quaisquer, os círculos de centro na origem e raios r_1 e r_2, isto é, os caminhos $\lambda_1, \lambda_2\colon [0, 2\pi] \to \mathbb{R}^2 - \{0\}$, $\lambda_1(s) = (r_1 \cos s, r_1 \operatorname{sen} s)$, $\lambda_2(s) = (r_2 \cos s, r_2 \operatorname{sen} s)$, são livrementre homotópicos, pela homotopia linear $H\colon [0, 2\pi] \times I \to \mathbb{R}^2 - \{0\}$, $H(s,t) = (1-t) \cdot \lambda_1(s) + t \cdot \lambda_2(s)$. Segue-se daí que, para toda forma fechada ω em $\mathbb{R}^2 - \{0\}$, vale $\int_{\lambda_1} \omega = \int_{\lambda_2} \omega$.

Corolário. *Seja $\lambda\colon [a,b] \to U$ um caminho fechado, seccionalmente C^1, livremente homotópico a um caminho constante no aberto $U \subset \mathbb{R}^n$. Para toda forma fechada $\omega\colon U \to (\mathbb{R}^n)^*$, tem-se $\int_\lambda \omega = 0$.*

Se quisermos provar que um caminho fechado $\lambda\colon [a,b] \to U$ *não* é livrementre homotópico a um caminho constante em U (fato que, em geral não é fácil de provar diretamente), basta obter uma 1-forma fechada $\omega\colon U \to (\mathbb{R}^n)^*$ tal que $\int_\lambda \omega \neq 0$. Por exemplo, o caminho fechado $\lambda\colon [0, 2\pi] \to \mathbb{R}^2 - \{0\}$, dado por $\lambda(t) = (\cos t, \operatorname{sen} t)$, não é livremente homotópico a um caminho constante em $\mathbb{R}^2 - \{0\}$ porque $\int_\lambda \delta\theta = 2\pi$.

Dizemos que um conjunto $X \subset \mathbb{R}^n$ é *simplesmente conexo* quando X é conexo por caminhos e, além disso, todo caminho fechado em X é livremente homotópico a um caminho constante.

Por exemplo, todo conjunto convexo $X \subset \mathbb{R}^n$ é simplesmente conexo pois se $\lambda\colon [a,b] \to X$ é um caminho fechado então $H\colon [a,b] \times [0,1] \to X$, dada por $H(s,t) = (1-t) \cdot \lambda(s) + t \cdot p$, onde $p = \lambda(a) = \lambda(b)$, é uma homotopia entre λ e o caminho constante p. Em particular, as bolas e os subespaços vetoriais de $\mathbb{R}^n$ são simplesmente conexos.

Por outro lado, $\mathbb{R}^2 - \{0\}$ não é simplesmente conexo pois o círculo $\lambda(t) = (\cos t, \operatorname{sen} t)$, $0 \leq t \leq 2\pi$ é um caminho fechado em $\mathbb{R}^2 - \{0\}$, o qual não é livremente homotópico a um caminho constante.

Observação: Dois caminhos fechados contidos em X, com o mesmo ponto básico podem ser livremente homotópicos em X sem serem ho-

motópicos em X (com ponto básico fixo). Há, porém, um caso particular importante em que homotopia livre implica homotopia com ponto básico fixo: se $\lambda\colon [a,b] \to X$ é um caminho fechado livremente homotópico a um caminho constante então λ é homotópico (com ponto básico fixo) a um caminho constante. (Para um estudo mais detalhado desta matéria, veja [10], pág. 47.) A definição usual de conjunto simplesmente conexo requer que todo caminho fechado seja homotópico (com ponto básico fixo) a um caminho constante. Adotamos uma definição equivalente, mais fácil de verificar.

Teorema 13. *Se o aberto $U \subset \mathbb{R}^n$ é simplesmente conexo então toda 1-forma fechada ω, em U é exata.*

Com efeito, para todo caminho fechado λ, seccionalmente C^1 em U, temos $\int_\lambda \omega = 0$. (Veja Teorema 9.)

Outra conseqüência interessante do teorema da invariância por homotopias (Teorema 11) é que se a forma $\omega\colon U \to (\mathbb{R}^n)^*$ é *fechada*, então a integral $\int_\lambda \omega$ tem sentido para qualquer caminho contínuo $\lambda\colon [a,b] \to U$, *mesmo que λ não seja retificável.* Com efeito, dado λ, existe $\lambda_1\colon [a,b] \to U$ poligonal (e portanto retificável) homotópico a λ (com extremidades fixas!).

Poremos então, por definição, $\int_\lambda \omega = \int_{\lambda_1} \omega$. Como λ_1 é contínuo e retificável, o segundo membro desta igualdade faz sentido. Sendo ω fechada, o valor $\int_{\lambda_1} \omega$ não depende de qual caminho poligonal λ_1, homotópico a λ, foi escolhido: se escolhêssemos λ_2, teríamos $\lambda \simeq \lambda_1$, $\lambda \simeq \lambda_2$, donde $\lambda_1 \simeq \lambda_2$ e portanto $\int_{\lambda_1} \omega = \int_{\lambda_2} \omega$.

Em particular se $\lambda\colon [a,b] \to \mathbb{R}^2 - \{0\}$ é qualquer caminho fechado (apenas contínuo), podemos definir sem ambigüidade o *número de voltas* que λ percorre em torno da origem como $n(\lambda) = (1/2\pi)\int_\lambda \delta\theta$, pois a forma elemento de ângulo $\delta\theta$ é fechada. Se λ e μ são caminhos fechados, livremente homotópicos em $\mathbb{R}^2 - \{0\}$, tem-se $n(\lambda) = n(\mu)$.

Os critérios aqui obtidos a fim de que uma forma seja exata se traduzem, analogamente, em condições para que um campo de vetores $F\colon U \to \mathbb{R}^n$, definido num aberto $U \subset \mathbb{R}^n$, seja um campo gradiente, isto é, que exista uma "função potencial" $f\colon U \to \mathbb{R}$, tal que $F(x) = \operatorname{grad} f(x)$ para todo $x \in U$.

Por exemplo, se o aberto $U \subset \mathbb{R}^n$ é simplesmente conexo, então o campo de vetores $F = (a_1, \ldots, a_n)\colon U \to \mathbb{R}^n$, de classe C^1, é o gradiente de um potencial se, e somente se, $\partial a_i/\partial x_j = \partial a_j/\partial x_i$ em todos os pontos de U, para quaisquer $i, j \in [1, n]$.

Dos fatos estabelecidos neste parágrafo resulta que se $U \subset \mathbb{R}^2$ é um aberto simplesmente conexo então:

1) Toda função harmônica $u\colon U \to \mathbb{R}$ é a parte real de uma função holomorfa $f\colon U \to \mathbb{C}$;

2) Toda função holomorfa $f\colon U \to \mathbb{C}$ possui uma primitiva $g\colon U \to \mathbb{C}$;

3) Se o aberto simplesmente conexo U não contém a origem 0, então existe uma função-ângulo contínua $\theta\colon U \to \mathbb{R}$.

Estas conclusões resultam do Teorema 2 e dos Exemplos 1 e 2 acima.

Como aplicação de 2), tomemos um aberto simplesmente conexo $U \subset \mathbb{R}^2 - \{0\}$. A função $f\colon U \to \mathbb{C}$, definida por $f(z) = 1/z$, tem uma primitiva $g\colon U \to \mathbb{C}$. A exponencial e^g cumpre: $(e^g/z)' = [g' \cdot e^g \cdot z - e^g]/z^2 = [e^g - e^g]/z^2 = 0$. Logo e^g/z é constante em U. Escrevendo essa constante sob a forma e^c, $c \in \mathbb{C}$, temos $e^{g(z)} = e^c \cdot z$ para todo $z \in U$. Então, pondo $h(z) = g(z) - c$, a função $h\colon U \to \mathbb{C}$ é holomorfa e tal que $e^{h(z)} = z$ para todo $z \in U$. Isto significa que h é um “ramo do logaritmo”, o qual portanto pode ser definido em qualquer aberto simplesmente conexo que não contenha a origem em $\mathbb{R}^2$.

Exemplo 3. *Integração complexa.* As integrais curvilíneas desempenham um papel fundamental na teoria das funções de uma variável complexa, como bem mostrou Cauchy. Dada uma função contínua $f = u + iv\colon U \to \mathbb{C}$, definida no aberto $U \subset \mathbb{C}$, para todo caminho contínuo retificável $\lambda\colon [a,b] \to U$ tem sentido a integral $\int_\lambda f(z)dz$. Aqui, $f(z)dz$ é uma forma diferencial complexa: $f(z)dz = (u+iv)(dx+idy) = udx - vdy + i(vdx + udy)$. Então, por definição, $\int_\lambda f(z)dz = \int_\lambda udx - vdy + i\int_\lambda vdx + udy$. Quando u e v são funções reais diferenciáveis em U, as equações de Cauchy-Riemann mostram que a forma $f(z)dz$ é fechada (isto é, sua parte real, $udx - vdy$, e sua parte imaginária $vdx + udy$ são ambas fechadas) se, e somente se, a função $f(z)$ é holomorfa. Então, quando $f\colon U \to \mathbb{C}$ é holomorfa (de classe C^1) e λ é um caminho fechado, livremente homotópico em U a um caminho constante, o Teorema 12 (ou mais precisamente, seu corolário) diz que $\int_\lambda f(z)dz = 0$. Esta é uma formulação do *Teorema de Cauchy*, uma das pedras fundamentais da bela teoria das funções complexas. Tomemos, como exemplo, $f\colon \mathbb{C} - \{0\} \to \mathbb{C}$,

$f(z) = 1/z$. Então

$$f(z)dz = \frac{dz}{z} = \frac{dx + idy}{x + iy} = \frac{(x - iy)(dx + idy)}{x^2 + y^2} = \\ = \frac{xdx + ydy + i(xdy - ydx)}{x^2 + y^2}.$$

Para todo caminho fechado $\lambda\colon [a, b] \to \mathbb{C} - \{0\}$, temos

$$\int_\lambda \frac{dz}{z} = \int_\lambda \frac{xdx + ydy}{x^2 + y^2} + i \int_\lambda \frac{xdy - ydx}{x^2 + y^2}.$$

Ora, a forma $(xdx + ydy)/(x^2 + y^2) = d(\log \sqrt{x^2 + y^2})$ é exata em $\mathbb{R}^2 - \{0\}$, logo a primeira integral no 2º membro da igualdade acima é zero.

O integrando da última integral acima é $\delta\theta$ logo

$$\frac{1}{2\pi i} \int_\lambda \frac{dz}{z} = \frac{1}{2\pi} \int_\lambda \frac{xdy - ydx}{x^2 + y^2} = \frac{1}{2\pi} \int_\lambda \delta\theta,$$

o que permite exprimir por meio de uma integral complexa o número de voltas que o caminho fechado λ percorre em torno da origem.

9 Cohomologia

Consideremos um aberto $U \subset \mathbb{R}^n$. Diremos que duas 1-formas fechadas ω_1, ω_2 em U, são *equivalentes* quando existir uma função $f\colon U \to \mathbb{R}$, de classe C^2, tal que $\omega_1 - \omega_2 = df$. Escreveremos então $\omega_1 \sim \omega_2$. Isto significa que, para todo caminho fechado λ, seccionalmente C^1 em U, tem-se $\int_\lambda \omega_1 = \int_\lambda \omega_2$.

A relação $\omega_1 \sim \omega_2$ é reflexiva, simétrica e transitiva; logo decompõe o conjunto das 1-formas fechadas em U como reunião disjunta de classes de equivalência. Para cada uma dessas formas ω, sua classe de equivalência é, por definição, o conjunto

$$[\omega] = \{\omega + df; f\colon U \to \mathbb{R} \text{ de classe } C^2\}.$$

Como se vê facilmente,

$$\omega_1 \sim \omega_2, \varphi_1 \sim \varphi_2 \Rightarrow \omega_1 + \varphi_1 \sim \omega_2 + \varphi_2 \Rightarrow [\omega_1 + \varphi_1] = \\ = [\omega_2 + \varphi_2] \text{ e } \omega_1 \sim \omega_2 \Rightarrow t \cdot \omega_1 \sim t \cdot \omega_2 \Rightarrow [t \cdot \omega_1] = [t \cdot \omega_2]$$

para todo número real t.

Isto quer dizer que as classes $[\omega + \varphi]$ e $[t \cdot \omega]$ dependem apenas das classes $[\omega]$ e $[\varphi]$ e do número real t, mas não das particulares formas ω, φ escolhidas para representá-las.

Podemos então introduzir a *soma* $[\omega]+[\varphi]$ de duas classes e o *produto* $t \cdot [\omega]$ de uma classe por um número real pondo, por definição:

$$[\omega] + [\varphi] = [\omega + \varphi],$$
$$t \cdot [\omega] = [t \cdot \omega].$$

Seja $H^1(U)$ o conjunto cujos elementos são as classes de equivalência $[\omega]$ das 1-formas fechadas em U. As operações acima tornam $H^1(U)$ um espaço vetorial real, tradicionalmente conhecido como o *grupo de cohomologia real de dimensão* 1 do aberto $U \subset \mathbb{R}^n$.

O elemento neutro (zero) do espaço vetorial $H^1(U)$ é a classe $[df]$ das 1-formas exatas em U.

A seguir, determinaremos a dimensão do (ou, mais precisamente, uma base para o) espaço vetorial $H^1(U)$ nos casos em que $U = \mathbb{R}^2 - \{0\}$, $U = \mathbb{R}^2 - \{p, q\}$ (p, q pontos distintos em $\mathbb{R}^2$) e $U = \mathbb{R}^n - \{0\}$, $n > 2$.

Inicialmente, seja $U = \mathbb{R}^2 - \{0\}$.

Neste caso, provaremos que $H^1(U)$ tem dimensão 1 e que podemos tomar como base a classe $[\delta\theta]$ da 1-forma elemento de ângulo.

Devemos então provar as duas afirmações seguintes:

1) $[\delta\theta] \neq 0$;

2) Para toda 1-forma fechada ω em $\mathbb{R}^2 - \{0\}$, existe um número real c tal que $[\omega] = c[\delta\theta]$.

A afirmação 1) significa que $\delta\theta$ não é uma forma exata em $\mathbb{R}^2 - \{0\}$, coisa que já sabemos.

A afirmação 2) corresponde a dizer que para toda 1-forma fechada ω em U, existe um número real c, tal que $\omega - c \cdot \delta\theta$ é exata em U. Isto é o que passamos a demonstrar. Utilizaremos o

Lema 1. *Seja* $\mu\colon [0, 2\pi] \to \mathbb{R}^2 - \{0\}$ *o caminho dado por* $\mu(s) = (\cos s, \operatorname{sen} s)$. *Para todo caminho fechado* $\lambda\colon [0, 2\pi] \to \mathbb{R}^2 - \{0\}$, *existe* $n \in \mathbb{Z}$ *tal que* λ *é livremente homotópico a* $n \cdot \mu$, *onde* $n \cdot \mu(s) = (\cos ns, \operatorname{sen} ns)$.

Demonstração: Pelo Lema 2 do §7, podemos supor que λ é de classe C^1 por partes. Para todo $s \in [0, 2\pi]$, o segmento de reta $\left[\lambda(s), \dfrac{\lambda(s)}{|\lambda(s)|}\right]$ não contém a origem. Logo λ é livremente homotópico em $\mathbb{R}^2 - \{0\}$, por uma homotopia linear, ao caminho $s \mapsto \dfrac{\lambda(s)}{|\lambda(s)|} \cdot$ Assim, não há perda de generalidade em supor que λ é um caminho de classe C^1 por partes, tal que $|\lambda(s)| = 1$ para todo $s \in [0, 2\pi]$. Seja então $\theta\colon [0, 2\pi] \to \mathbb{R}$ uma função-ângulo para λ, de modo que $\lambda(s) = (\cos\theta(s), \operatorname{sen}\theta(s))$ para todo $s \in [0, 2\pi]$). (Veja o Corolário 2 do Teorema 11, Capítulo II.) Como $\lambda(0) = \lambda(2\pi)$, existe $n \in \mathbb{Z}$ tal que $\theta(2\pi) = \theta(0) + 2n\pi$. A função contínua $\widetilde{H}\colon [0, 2\pi] \times [0, 1] \to \mathbb{R}$ definida por $\widetilde{H}(s, t) = (1 - t)\theta(s) + t \cdot n \cdot s$, cumpre $\widetilde{H}(2\pi, t) = \widetilde{H}(0, t) + 2n \cdot \pi$ para todo $t \in [0, 1]$. Segue-se daí que a aplicação $H\colon [0, 2\pi] \times [0, 1] \to \mathbb{R}^2 - \{0\}$, definida por $H(s, t) = (\cos\widetilde{H}(s, t), \operatorname{sen}\widetilde{H}(s, t))$, é uma homotopia livre entre os caminhos fechados λ e $n \cdot \mu$.

O cálculo do grupo de cohomologia $H^1(\mathbb{R}^2 - \{0\})$ se faz então com o teorema abaixo. Antes de prová-lo, observemos que se $n \cdot \mu$ é o caminho definido no enunciado do Lema 1, então $n \cdot \mu = \mu + \mu + \cdots + \mu$ (n parcelas), de modo que $\int_{n \cdot \mu} \omega = n \cdot \int_\mu \omega$ para toda forma contínua ω, definida num conjunto que contenha o círculo S^1, imagem de μ. Aqui, evidentemente $\mu = 1 \cdot \mu$ é o caminho $\mu(s) = (\cos s, \operatorname{sen} s)$.

Teorema 14. *Para toda 1-forma fechada ω em $\mathbb{R}^2 - \{0\}$, existe um número real c tal que a 1-forma $\omega - c \cdot \delta\theta$ é exata em U.*

Demonstração: Ponha $c = \dfrac{1}{2\pi} \displaystyle\int_\mu \omega$, onde $\mu(s) = (\cos s, \operatorname{sen} s)$, $s \in [0, 2\pi]$ é o caminho circular do Lema 1. Devemos provar que $\int_\lambda (\omega - c \cdot \delta\theta) = 0$, ou ainda, que $\int_\lambda \omega = c \cdot \int_\lambda \delta\theta$ para todo caminho fechado λ em $\mathbb{R}^2 - \{0\}$. Reparametrizando λ, se necessário, podemos supô-lo definido no intervalo $[0, 2\pi]$, pois $\int_\lambda \omega$ é invariante por reparametrizações positivas. Pelo Lema 1, λ é livremente homotópico a $n \cdot \mu$ para algum $n \in \mathbb{Z}$. Como ω e $\delta\theta$ são fechadas, suas integrais sobre caminhos fechados livremente homotópicos são iguais. Então

$$\int_\lambda \omega = \int_{n\mu} \omega = n \int_\mu \omega = n \cdot 2\pi c = c \cdot n \int_\mu \delta\theta = c \int_{n\mu} \delta\theta = c \cdot \int_\lambda \delta\theta,$$

como queríamos demonstrar.

Em seguida, mostraremos que se $p, q \in \mathbb{R}^2$ são pontos distintos então o espaço vetorial $H^1(\mathbb{R}^2 - \{p, q\})$ tem dimensão 2.

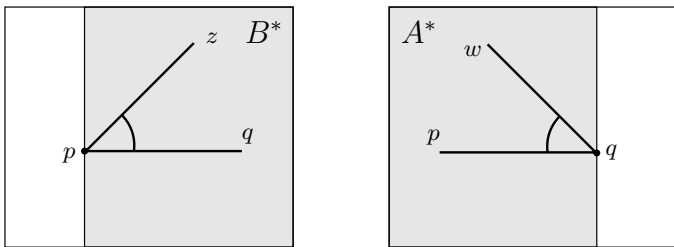
Começaremos introduzindo, para cada ponto $p = (a, b) \in \mathbb{R}^2$, a 1-forma fechada α_p definida em $\mathbb{R}^2 - \{p\}$, chamada o "elemento de ângulo com vértice p". Dado $z = (x, y) \in \mathbb{R}^2 - \{p\}$, poremos:

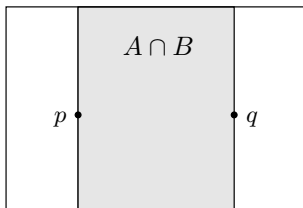
$$\alpha_p(z) = \frac{1}{|z-p|^2}\,[-(y-b)dx + (x-a)dy].$$

Seja $\mu_p \colon [0, 2\pi] \to \mathbb{R}^2 - \{p\}$ o caminho circular de centro p (e raio r, definido por $\mu_p(s) = p + r \cdot e^{is}$. Então $\int_{\mu_p} \alpha_p = 2\pi$.

Por outro lado, se $q \neq p$ e μ_q é um caminho circular de centro q e raio $r < |p-q|$ então, tomando $\varepsilon > 0$ tal que $r < r + \varepsilon < |q-p|$, o disco aberto de centro q e raio $r + \varepsilon$ é um aberto simplesmente conexo que não contém p, logo a forma α_p está definida nele e é, portanto, exata ali. Segue-se que $\int_{\mu_q} \alpha_p = 0$ quando o raio r do círculo μ_q é inferior à distância $|q-p|$.

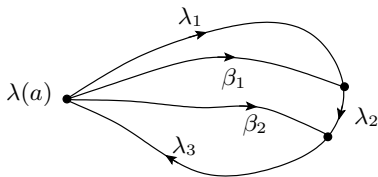
Fixemos agora, definitivamente, dois pontos $p \neq q$ no plano $\mathbb{R}^2$. Sejam

$$B^* = \{z \in \mathbb{R}^2; \langle z-p, q-p\rangle > 0\},\ A^* = \{w \in \mathbb{R}^2; \langle w-q, p-q\rangle > 0\}.$$

Temos $z \in B^* \Leftrightarrow$ ângulo $\widehat{zpq}$ é agudo, enquanto $w \in A^* \Leftrightarrow$ o ângulo $\widehat{wqp}$ é agudo.

Ponhamos agora $A = A^* - \{p\}$, $\;B = B^* - \{q\}$. Então A e B são abertos, $\mathbb{R}^2 - \{p, q\} = A \cup B$ e $A \cap B$ é convexo.

Teorema 15. *Dada uma* 1*-forma fechada* ω *em* $\mathbb{R}^2 - \{p, q\}$*, existem números reais univocamente determinados* u*,* v *tais que a* 1*-forma* $\varphi = \omega - u \cdot \alpha_p - v \cdot \alpha_q$ *é exata em* $\mathbb{R}^2 - \{p, q\}$*. (Isto é,* $[\omega] = u \cdot [\alpha_p] + v \cdot [\alpha_q]$*.)*

Demonstração: Sejam μ_p um caminho circular em A, com centro p e μ_q um caminho circular em B, de centro q. Ponhamos $u = \dfrac{1}{2\pi}\displaystyle\int_{\mu_p} \omega$, $v = \dfrac{1}{2\pi}\displaystyle\int_{\mu_q} \omega$ e $\varphi = \omega - u\alpha_p - v\alpha_q$. Como no Lema 1, vê-se que todo caminho fechado em A é livremente homotópico a um múltiplo inteiro de μ_p. Como $\int_{\mu_p} \varphi = 0$, segue-se que $\int_\lambda \varphi = 0$ para todo caminho fechado λ em A. Assim φ é exata em A. Analogamente, concluimos que φ é exata em B. A esta altura, faremos uso do

Lema 2. *Sejam* $A, B \subset \mathbb{R}^n$ *abertos conexos, com* $A \cap B$ *conexo, e* $\varphi\colon A \cup B \to (\mathbb{R}^n)^*$ *uma* 1*-forma fechada. Se* φ *é exata em* A *e em* B *separadamente, então* φ *é exata em* $A \cup B$*.*

Com efeito, sabemos que $\int_\lambda \varphi = 0$ para todo caminho fechado λ contido em A ou contido em B. Devemos provar que $\int_\lambda \varphi = 0$ para todo caminho fechado $\lambda\colon [a, b] \to A \cup B$. Como A e B são abertos, existe uma partição $P = \{a = t_0 < t_1 < \cdots < t_k = b\}$ tal que, para todo $i = 1, \ldots, k$, a restrição $\lambda_i = \lambda|[t_{i-1}, t_i]$ é um caminho contido em A ou contido em B. Eliminando pontos supérfluos t_i na partição P, podemos supor que λ_i e λ_{i+1} não estão ambos em A nem ambos em B. Então $\lambda(t_i) \in A \cap B$ para todo $i = 1, 2, \ldots, k-1$. Suporemos inicialmente que $\lambda(a) = \lambda(b) \in A \cap B$. Como $A \cap B$ é conexo, existem caminhos $\beta_1, \ldots, \beta_{k-1}$ em $A \cap B$, cada β_i ligando o ponto $\lambda(a)$ ao ponto $\lambda(t_i)$. Note que $\lambda = \lambda_1 + \cdots + \lambda_k$ e que os caminhos $\lambda_1 - \beta_1$, $\beta_1 + \lambda_2 - \beta_2, \ldots, \beta_{k-2} + \lambda_{k-1} - \beta_{k-1}$, $\beta_{k-1} + \lambda_k$ são fechados, cada um deles inteiramente contidos em A ou em B.

Logo, a integral de φ ao longo de cada um destes caminhos é zero.

Portanto

$$\int_\lambda \omega = \sum_{i=1}^k \int_{\lambda_i} \omega = \int_{\lambda_1} \omega - \int_{\beta_1} \omega + \int_{\beta_1} \omega + \int_{\lambda_2} \omega + \cdots + \\ + \int_{\beta_{k-1}} \omega + \int_{\lambda_k} \omega = \int_{\lambda_1 - \beta_1} \omega + \int_{\beta_1 + \lambda_2 - \beta_2} \omega + \cdots + \\ + \int_{\beta_{k-1} + \lambda_k} \omega = 0.$$

Se, por acaso, não for $\lambda(a) = \lambda(b) \in A \cap B$, tomamos um caminho β em $A \cup B$ ligando um ponto de $A \cap B$ ao ponto $\lambda(a) = \lambda(b)$. Então o caminho $\beta + \lambda - \beta = \gamma$ começa e termina em $A \cap B$ e temos $\int_\lambda \omega = \int_\gamma \omega = 0$.

Resta apenas mostrar que a unicidade dos coeficientes u, v, ou, o que é o mesmo, a independência linear das classes $[\alpha_p]$ e $[\alpha_q]$ em $H^1(\mathbb{R}^2 - \{p,q\})$. Ora, de $\omega = u \cdot \alpha_p + v \cdot \alpha_q + df$, obtemos $(1/2\pi) \int_\lambda \omega = u$ para todo caminho circular λ, de centro p, que não contenha q em seu interior. Isto mostra que u é univocamente determinado a partir de ω. Do mesmo modo se raciocina com respeito a v.

Observação: Este método serve para provar, mais geralmente, que se $p_1, \ldots, p_n$ são n pontos distintos em $\mathbb{R}^2$ e $U = \mathbb{R}^2 - \{p_1, \ldots, p_n\}$ então $H^1(U)$ é um espaço vetorial de dimensão n.

Mostraremos, finalmente, que se $U = \mathbb{R}^n - \{0\}$ com $n \geq 3$ então $H^1(U) = \{0\}$, isto é, que toda 1-forma fechada em U é exata.

Para isto, é suficiente provar que, se $n > 2$, todo caminho fechado em $\mathbb{R}^n - \{0\}$ é livremente homotópico a um caminho constante.

Em virtude do Lema 2, §7, podemos partir de um caminho poligonal λ em $\mathbb{R}^n - \{0\}$.

Se existirem três vértices consecutivos A, B C do caminho poligonal λ tais que o triângulo ABC não contenha a origem $0 \in \mathbb{R}^n$, então podemos, através de uma homotopia linear em $\mathbb{R}^n - \{0\}$, deformar $AB + BC$ em AC, logo λ será homotópico em U a um caminho poligonal fechado, com um vértice a menos. Repetindo o processo, eliminaremos um a um todos os vértices, chegando a um caminho, que ainda chamaremos de λ, tal que qualquer triângulo determinado por 3 vértices consecutivos de λ contém a origem. Então o caminho λ está contido num plano Π, um subespaço de dimensão 2 em $\mathbb{R}^n$. Como $n \geq 3$, existe um vetor $a \in \mathbb{R}^n$ que não pertence a esse plano Π. Então $H(s,t) = \lambda(s) + ta$ é uma homotopia

em $\mathbb{R}^n - \{0\}$, começando com λ e terminando no caminho $s \mapsto \lambda(s) + a$, contido no plano $\Pi + a$ e, neste plano, homotópico livremente a um caminho constante. Como $0 \notin \Pi + a$, esta última homotopia tem lugar em $\mathbb{R}^n - \{0\}$, o que termina a demonstração.

Acabamos de ver que, se $n \geq 2$, o conjunto $\mathbb{R}^{n+1} - \{0\}$ é simplesmente conexo. Isto tem a seguinte conseqüência:

Para $n \geq 2$, a esfera n-dimensional S^n é simplesmente conexa.

De fato, como $S^n \subset \mathbb{R}^{n+1} - \{0\}$, todo caminho fechado λ em S^n é livremente homotópico em $\mathbb{R}^{n+1} - \{0\}$ a um caminho constante. Compondo essa homotopia com a "projeção radial" $\rho\colon \mathbb{R}^{n+1} - \{0\} \to S^n$, $\rho(x) = \dfrac{x}{|x|}$, obtemos uma homotopia livre em S^n, entre o caminho fechado λ e um caminho constante.

10 A fórmula de Kronecker

Uma aplicação interessante das integrais curvilíneas é a fórmula de Kronecker, sobre o número de raízes de um sistema de equações do tipo $f(x,y) = 0$, $g(x,y) = 0$, onde f e g são funções de classe C^2 definidas num disco do plano.

Começaremos com algumas considerações gerais sobre mudança de variáveis em formas diferenciais.

Diremos que uma aplicação $F = (f_1, \ldots, f_n)\colon U \to \mathbb{R}^n$, definida num aberto $U \subset \mathbb{R}^m$, *é de classe C^k* quando cada uma das suas funções coordenadas $f_k\colon U \to \mathbb{R}$ $(1 \leq k \leq n)$ for de classe C^k. Analogamente se define aplicação diferenciável $F\colon U \to \mathbb{R}^n$. Essas aplicações serão objeto de estudo detalhado no capítulo seguinte.

Dadas uma forma diferencial $\omega = \sum\limits_{k=1}^{n} a_k dy_k$, de classe C^r, definida no aberto $V \subset \mathbb{R}^n$, e uma aplicação $F = (f_1, \ldots, f_n)\colon U \to V$, de classe C^{r+1} no aberto $U \subset \mathbb{R}^m$, definiremos a *forma induzida* $F^*\omega$ em U pondo

$$F^*\omega = \sum_{k=1}^{n} (a_k \circ F) df_k \,.$$

Para todo $x \in U$ temos, portanto

$$(F^*\omega)(x) = \sum_{i=1}^{m} b_i(x) dx_i \,, \text{ onde } b_i(x) = \sum_{k=1}^{n} a_k(F(x)) \frac{\partial f_k}{\partial x_i}(x).$$

Assim, a forma induzida $F^*\omega$ é obtida de ω pela "mudança de variáveis" $y = F(x)$, isto é, fazendo-se as substituições $y_k = f_k(x)$, $dy_k = df_k = \Sigma \dfrac{\partial f_k}{\partial x_i} dx_i$.

Como se vê, para $\omega \in C^r$ e $F \in C^{r+1}$, tem-se $F^*\omega \in C^r$.

As propriedades seguintes são de verificação imediata:

1. Se $F\colon U \to V$ e $G\colon V \to W$ de classe C^{r+1} então, para toda forma ω de classe C^r em W, vale $(G \circ F)^*\omega = F^*(G^*\omega)$;

2. Se $\omega = df$ é exata, então $F^*\omega = d(f \circ F)$ também é exata;

3. Se ω é fechada, então $F^*\omega$ também é fechada;

4. Para todo caminho contínuo retificável λ em U, $\int_\lambda F^*\omega = \int_{F\circ\lambda} \omega$.

Um ponto $x \in U \subset \mathbb{R}^n$ chama-se um *zero* da aplicação $F\colon U \to \mathbb{R}^n$ quando $F(x) = 0$. Dizemos que x é um *zero isolado* de F quando existe uma bola $B = B(x, r) \subset U$ tal que x é o único zero de F em B. O zero x chama-se *simples* quando F é diferenciável e o *determinante Jacobiano* $JF(x) = \det\left(\dfrac{\partial f_i}{\partial x_j}(x)\right)$ é diferente de zero. Se for $JF(x) > 0$, diremos que x é um *zero positivo* de F. Analogamente, diremos que x é um *zero negativo* de F quando tivermos $JF(x) < 0$.

Pelo Teorema 4a do Capítulo III, se F é diferenciável então todo zero simples de F é isolado.

Passemos agora ao caso particular que nos interessa no momento.

Dado um círculo C, de centro p e raio r no plano, indicaremos ainda com C o caminho $C\colon [0, 2\pi] \to \mathbb{R}^2$, definido por $C(s) = p + r \cdot e^{is}$.

Sejam $F = (f, g)\colon U \to \mathbb{R}^2$ uma aplicação de classe C^2 no aberto $U \subset \mathbb{R}^2$ e $D \subset U$ um disco fechado, em cuja fronteira $C = \partial D$ não há zeros de F. A integral

$$n(F, D) = \frac{1}{2\pi} \int_C \frac{f dg - g df}{f^2 + g^2}$$

chama-se o *índice* da aplicação F no disco D.

A forma diferencial $(f dg - g df)/(f^2 + g^2)$ acha-se definida no aberto V, formado pelos pontos $z \in U$ tais que $F(z) \neq 0$ (o que equivale a $f(z)^2 + g(z)^2 > 0$). Em V, esta forma é igual a $F^*(\delta\theta)$. Por hipótese, temos $C \subset V$. Então

$$n(f; D) = \frac{1}{2\pi} \int_C \frac{f dg - g df}{f^2 + g^2} = \frac{1}{2\pi} \int_C F^*(\delta\theta) = \frac{1}{2\pi} \int_{F\circ C} \delta\theta.$$

Em particular, o índice $n(F;D)$ é um número inteiro, igual ao número de voltas que o caminho $F \circ C$ percorre em torno da origem em $\mathbb{R}^2$. (Veja o Corolário do Teorema 8.)

Teorema 16. (1) *Se $D_1 \subset D_2$ são discos contidos no aberto $U \subset \mathbb{R}^2$ e $F\colon U \to \mathbb{R}^2$ não possui zeros no conjunto $D_2 - \operatorname{int} D_1$ então $n(F;D_1) = n(F;D_2)$.*

(2) *Sejam $F, G\colon U \to \mathbb{R}^2$ aplicações de classe C^2, que não possuem zeros na fronteira $C = \partial D$ de um disco D contido em U. Suponha que exista uma homotopia $H\colon C \times [0,1] \to \mathbb{R}^2 - \{0\}$ entre $F|C$ e $G|C$ (isto é, $H(z,0) = F(z)$, $H(z,1) = G(z)$ e $H(z,t) \neq 0$, para todo $z \in C$ e todo $t \in [0,1]$). Então $n(F;D) = n(G;D)$.*

Em particular, se $0 \notin [F(z), G(z)]$ para todo $z \in C$ então $n(F;D) = n(;D)$.

Demonstração: (1) Sejam $C_1 = \partial D_1$ e $C_2 = \partial D_2$ parametrizados por $C_1(s) = p_1 + r_1 \cdot e^{is}$ e $C_2(s) = p_2 + r_2 \cdot e^{is}$, $0 \leq s \leq 2\pi$. Então $H\colon [0,2\pi]\times[0,1] \to \mathbb{R}^2 - \{0\}$, dada por $H(s,t) = F[(1-t)C_1(s) + tC_2(s)]$, é uma homotopia livre entre os caminhos fechados $F \circ C_1$ e $F \circ C_2$. Como $\delta\theta$ é uma forma fechada em $\mathbb{R}^2 - \{0\}$, segue-se do Teorema 11 que

$$n(F;D_1) = \frac{1}{2\pi}\int_{F\circ C_1} \delta\theta = \frac{1}{2\pi}\int_{F\circ C_2} \delta\theta = n(F;D_2).$$

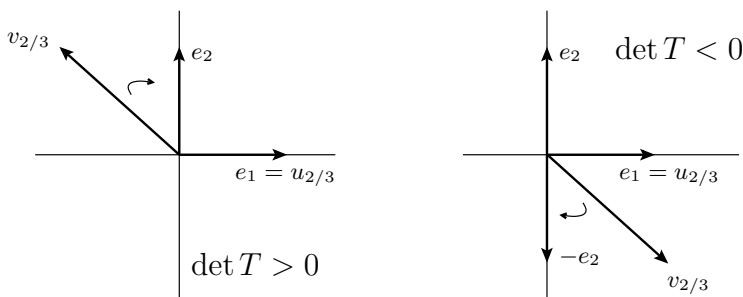
(2) A aplicação $K\colon [0,2\pi] \times [0,1] \to \mathbb{R}^2 - \{0\}$, definida por $K(s,t) = H(C(s),t)$, é uma homotopia livre entre os caminhos fechados $F \circ C$ e $G \circ C$. Logo, pelo Teorema 11,

$$n(F;D) = \frac{1}{2\pi}\int_{F\circ C} \delta\theta = \frac{1}{2\pi}\int_{G\circ C} \delta\theta = n(G;D).$$

Corolário. *Se $F\colon U \to \mathbb{R}^2$ (uma aplicação de classe C^2) não possui zeros no disco $D \subset U$ então $n(F;D) = 0$.*

O corolário acima pode ser enunciado, equivalentemente, sob a forma de um importante teorema de existência, assim: *Se $n(F;D) \neq 0$ então existe algum $z \in D$ tal que $F(z) = 0$.*

Com efeito, sendo p o centro do disco D, a aplicação $H\colon [0,2\pi] \times [0,1] \to \mathbb{R}^2 - \{0\}$, definida por $H(s,t) = F((1-t)\cdot C(s) + t\cdot p)$, é uma homotopia livre entre o caminho fechado $F \circ C$ e o caminho constante $F(p)$. Logo $n(F;D) = \dfrac{1}{2\pi}\displaystyle\int_{F\circ C} \delta\theta = \dfrac{1}{2\pi}\displaystyle\int_{F(p)} \delta\theta = 0.$

Exemplo 4. Seja $T: \mathbb{R}^2 \to \mathbb{R}^2$ uma transformação linear invertível. Se o disco $D \subset \mathbb{R}^2$ não contém a origem então T não possui zeros em D, logo $n(T;D) = 0$. Se D contém a origem em seu interior, provaremos que $n(T;D) = +1$ ou $n(T;D) = -1$, conforme seja $\det T > 0$ ou $\det T < 0$. Para isto mostraremos que existe uma homotopia entre T e uma transformação linear L que, no caso de $\det T > 0$, é a identidade e, quando $\det T < 0$, é a transformação $T_1(x,y) = (x,-y)$. Tal homotopia consiste num caminho contínuo $t \mapsto T_t$, $t \in [0,1]$, com $T_0 = T$ e $T_1 = L$, tal que, para todo $t \in [0,1]$, $T_t: \mathbb{R}^2 \to \mathbb{R}^2$ é uma transformação linear *invertível*. Então $n(T;D) = n(L;D)$, sendo óbvio que $n(L;D) = \pm 1$. Para obter o caminho de transformações lineares invertíveis T_t ligando T a L, trabalharemos com a matriz $\binom{ab}{cd}$ de T, a qual tem como colunas os vetores $u = (a,c) = T \cdot e_1$ e $v = (b,d) = T \cdot e_2$. Devemos obter um caminho contínuo $t \mapsto (u_t, v_t)$ tal que $u_0 = u$, $v_0 = v$, e $u_1 = e_1$, $v_1 = \pm e_2$ (conforme o sinal do determinante de T), de modo que, para todo t, os vetores u_t, v_t sejam linearmente independentes. Começamos deformando continuamente os vetores u, v ao longo das retas que os contêm, de modo a ficar com $|u| = |v| = 1$. A segunda parte do caminho (u_t, v_t) consiste em girar u no sentido horário, até fazê-lo coincidir com e_1, e girar v ao mesmo tempo, do modo a mantê-lo rigidamente ligado a u. Digamos que a primeira deformação levou $1/3$ da unidade de tempo para ser efetuada e

a segunda gastou outro $1/3$. Então, para $t = 2/3$ temos $u_{2/3} = e_1$ e $v_{2/3}$ pode estar (1) no semi-plano superior ou (2) no semi-plano inferior. Durante a deformação contínua dos vetores u_t, v_t, que são as colunas da matriz de T_t, o determinante de T_t varia continuamente e nunca se anula; logo mantém o mesmo sinal. Então o caso (1) corresponde a $\det T > 0$ e o caso (2) corresponde a $\det T < 0$. Para completar a homotopia, usamos o intervalo de tempo $[2/3, 1]$ para girar continuamente $v_{2/3}$, mantendo

$u_{2/3} = e_1$ fixo, até obter no caso (1), $v_1 = e_2$ e, no caso (2), obtendo $v_1 = -e_2$, isto sendo feito de modo que v_t e $u_t = e_1$ sejam linearmente independentes para todo $t \in [2/3, 1]$, o que é obviamente possível.

Teorema 17. *Se $F = (f, g)\colon U \to \mathbb{R}^2$, de classe C^2 no aberto $U \subset \mathbb{R}^2$, possui um único zero p no disco $D \subset U$, o qual é simples e não está em $C = \partial D$, então $n(F; D) = \pm 1$, conforme o Jacobiano $JF(p) = \dfrac{\partial f}{\partial x}\dfrac{\partial g}{\partial y} - \dfrac{\partial f}{\partial y}\dfrac{\partial g}{\partial x}$, no ponto p, seja positivo ou negativo.*

Demonstração: Para simplificar a notação, suporemos que p é a origem de $\mathbb{R}^2$, escreveremos $a = \dfrac{\partial f}{\partial x}(p)$, $b = \dfrac{\partial f}{\partial y}(p)$, $c = \dfrac{\partial g}{\partial x}(p)$, $d = \dfrac{\partial g}{\partial y}(p)$ e indicaremos com $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ a transformação linear (invertível) cuja matriz tem as linhas (a, b) e (c, d), nesta ordem. Como f e g são diferenciáveis, para todo $z = (x, y) \in D$ temos

$$f(z) = ax + by + \rho(z) \cdot |z| \ \text{ e } \ g(z) = cx + dy + \sigma(z) \cdot |z|,$$

$$\text{onde} \quad \lim_{z \to 0} \rho(z) = \lim_{z \to 0} \sigma(z) = 0.$$

Pondo $R(z) = (\rho(z), \sigma(z))$, podemos então escrever, para todo $z \in D$:

$$F(z) = T \cdot z + R(z) \cdot |z|, \quad \text{onde} \quad \lim_{z \to 0} R(z) = 0.$$

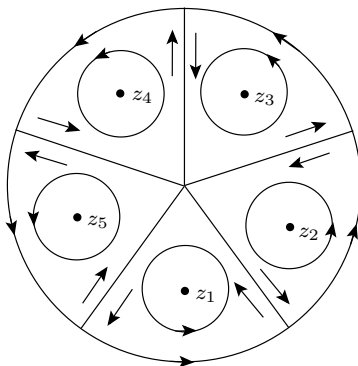
Sento T invertível, o Lema do Teorema 4a, Capítulo III, nos fornece $c > 0$ tal que $|T \cdot z| \geq c|z|$ para todo $z \in \mathbb{R}^2$. Pelo Teorema 16, podemos supor o raio do disco D tão pequeno que $z \in D \Rightarrow |R(z)| \leq c/2$. Então

$$0 \neq z \in D \Rightarrow |T \cdot z + (1-t) \cdot R(z)|z| \geq$$
$$|T \cdot z| - (1-t) \cdot |R(z)| \cdot |z| \geq c|z| - \frac{c}{2}\,|z| > 0,$$

para todo $t \in [0, 1]$. Segue-se que $H\colon U \times [0, 1] \to \mathbb{R}^2$, definida por $H(z, t) = T \cdot z + (1-t) \cdot R(z) \cdot |z|$, é uma homotopia entre F e T, tal que $H(z, t) \neq 0$ para todo $z \in C = \partial D$ e todo $t \in [0, 1]$. Portanto, $n(F; D) = n(T; D)$ em virtude do item (2) do Teorema 16. Finalmente, o Exemplo 4 nos dá $n(T; D) = \pm 1$, onde o sinal é o do Jacobiano de F no ponto p.

Teorema 18. *Se a aplicação $F = (f, g)\colon U \to \mathbb{R}^2$ é de classe C^2 e possui, no disco $D \subset U$, um número finito de zeros, todos simples, nenhum dos quais em $C = \partial D$, então $n(F; D) = P - N$, onde P é o*

número de zeros positivos de F em D e N é o número de zeros negativos. Noutros termos, temos a igualdade

$$\frac{1}{2\pi}\int_c \frac{f dg - g df}{f^2+g^2} = P - N.$$

Demonstraçao: Sejam $z_1, \ldots, z_k$ os zeros de F em D. Modificando ligeiramente o disco D, se necessário [segundo o item (1) do Teorema 16], podemos supor que nenhum raio de D contém mais de um zero de F. Por meio de raios, dividimos D em k setores S_i, cada um deles contendo exatamente um dos pontos z_i em seu interior. A integral $\int_c (f dg - g df)/(f^2+g^2)$ é a soma de k integrais com o mesmo integrando, cada uma delas tomada ao longo de ∂S_i, cada uma dessas fronteiras parametrizada de modo coerente com o arco de C que lhe é comum. Com centro em cada zero z_i, tomemos um pequeno círculo $C_i = \partial D_i$, contido no interior do setor S_i. Uma homotopia radial evidente entre C_i e ∂S_i nos dá $n(F; D) = \sum_{i=1}^{k} n(F; D_i)$ e o resultado então segue-se do Teorema 17.

Exemplo 5. Seja $F = u + i \cdot v \colon U \to \mathbb{C}$ uma função holomorfa. Em virtude das equações de Cauchy-Riemann, em todo ponto $z \in U$, seu determinante Jacobiano é $JF(z) = \frac{\partial u}{\partial x}\frac{\partial v}{\partial y} - \frac{\partial u}{\partial y}\frac{\partial v}{\partial x} = \left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 = |F'(z)|^2$. Assim, $JF(z) \geq 0$ para todo $z \in U$. Em particular, se todos os zeros de F no disco $D \subset U$ forem simples, então o índice $n(F; D)$ é igual ao número desses zeros. Consideremos o caso especial de um polinômio $p \colon \mathbb{C} \to \mathbb{C}$. Um zero $a \in \mathbb{C}$ chama-se então uma *raiz* de p. Como se sabe,

$p(a) = 0$ equivale a $p(z) = (z-a)\cdot q(z)$, onde $q(z)$ é ainda um polinômio. A *multiplicidade* da raiz a é o maior número natural k tal que p é divisível por $(z-a)^k$. [Isto equivale a dizer que $p(z) = (z-a)^k \cdot q(z)$, com $q(a) \neq 0$.] Quando $k = 1$, a diz-se uma *raiz simples*. Ora, $p(z) = (z-a)\cdot q(z)$ implica $p'(a) = q(a)$. Como um polinômio é uma função holomorfa, seu Jacobiano no ponto a é $Jp(a) = |p'(a)|^2 = |q(a)|^2$. Assim, a é um zero simples de p se, e somente se, é uma raiz simples. Suponhamos que o polinômio p só tenha raízes simples. Então, para todo círculo $C = \partial D$ no plano, sobre o qual não haja raízes de p, o número $n(p; D)$, ou seja, o número de voltas que o caminho $p \circ C$ descreve em torno da origem, é igual ao número de raízes de p no interior do círculo C (isto é, no disco D). Estendamos este resultado para um polinômio qualquer.

Se a é uma raiz, de multiplicidade k, do polinômio p, então $p(z) = (z-a)^k[c + \varphi(z)]$, onde $c \neq 0$ e $\varphi(z)$ é um polinômio, com $\varphi(a) = 0$. Existe portanto $\delta > 0$ tal que

$$0 < |z-a| \leq \delta \Rightarrow |\varphi(z)| < \frac{|c|}{3} \Rightarrow c + (1-t)\varphi(z) \neq 0$$
$$\text{para todo } t \in [0,1].$$

Sejam $D = B[a;\delta]$ e $C = \partial D$. Então $H\colon C \times [0,1] \to \mathbb{C}$, definida por $H(z,t) = (z-a)^k[c+(1-t)\varphi(z)]$, é uma homotopia entre $H(z,0) = p(z)$ e $H(z,1) = c(z-a)^k$, com a propriedade de que $H(z,t) \neq 0$ para todo $z \in C$ e todo $t \in [0,1]$. Concluimos, pelo Teorema 16, que $n(p;D) = n(c(z-a)^k; D) =$ número de voltas que o caminho fechado $s \mapsto \delta e^{iks}$, $0 \leq s \leq 2\pi$, percorre em torno da origem. Evidentemente, este número é k. Logo $n(p;D) = k$, sempre que o polinômio p tenha no disco D uma única raiz, $a \in \operatorname{int} D$, de multiplicidade k. Segue-se daí (como na demonstração do Teorema 18) que para todo disco $D \subset \mathbb{R}^2$ que não contenha raízes de p em sua fronteira C, o número $n(p;D)$, igual ao número de voltas que o caminho $p \circ C$ percorre em torno da origem em $\mathbb{C}$, é também igual ao número de raízes de p (cada raiz de multiplicidade k sendo contada k vezes) contidas no disco D. Ora, se o raio do círculo C for suficientemente grande, o polinômio p tiver grau m então o caminho fechado $p \circ C$ dá m voltas em torno da origem. Com efeito, se $p(z) = a_0 z^m + a_1 z^{m-1} + \cdots + a_m$ então

$$z \neq 0 \Rightarrow p(z) = z^m \left[a_0 + \frac{a_1}{z} + \cdots + \frac{a_m}{z^m}\right] = z^m[a_0 + \xi(z)],$$

onde $a_0 \neq 0$ e $\lim_{|z|\to\infty} \xi(z) = 0$. Existe portanto $r > 0$ tal que

$$|z| \geq r \Rightarrow |\xi(z)| < \frac{|a_0|}{2} \Rightarrow a_0 + (1-t)\,\xi(z) \neq 0 \text{ para todo } t \in [0,1].$$

Então $H\colon \mathbb{C}\times[0,1]\to\mathbb{C}$, definida por $H(z,t)=z^m[a_0+(1-t)\xi(z)]$, é uma homotopia entre $H(z,0) = p(z)$ e $H(z,1) = a_0 z^m$, com a propriedade de que $H(z,t) \neq 0$ para todo $t \in [0,1]$ e $|z| \geq r$. Se $D = B[0;r]$, temos então $n(p;D) = n(a_0 z^m; D) = m$. Em particular obtemos o Teorema Fundamental da Álgebra: todo polinômio complexo de grau ≥ 1 tem pelo menos uma raiz complexa. (Veja o corolário do Teorema 16.)

Como aplicação do Exemplo 5 acima, demonstraremos que as raízes de um polinômio dependem continuamente dos seus coeficientes. Mais precisamente, temos o

Teorema 19. *Sejam $p(z) = a_0 z^m + a_1 z^{m-1} + \cdots + a_m$ um polinômio complexo, $a \in \mathbb{C}$ uma raiz de p com multiplicidade k e D um disco de centro a e raio ε, não contendo outra raiz de p. Existe $\delta > 0$ tal que se $q(z) = b_0 z^m + b_1 z^{m-1} + \cdots + b_m$ é qualquer polinômio satisfazendo $|b_0 - a_0| < \delta, \ldots, |b_m - a_m| < \delta$ então q possui k raízes no disco D (cada uma delas sendo contada de acordo com sua multiplicidade).*

Demonstração: Como vimos no Exemplo 5, $k = n(p;D)$. O círculo $C = \partial D$ evidentemente é um conjunto limitado: $z \in C \Rightarrow |z| \leq |a| + \varepsilon$. Ponhamos $A = \sum_{i=0}^{m} (|a| + \varepsilon)^{m-i}$. Seja $\eta > 0$ o ínfimo de $|p(z)|$ quando z percorre o conjunto C. Assim, $z \in C \Rightarrow |p(z)| \geq \eta$. Tomemos $\delta = \eta/2A$. Se $q(z) = b_0 z^m + \cdots + b_m$ é tal que $|b_i - a_i| < \delta$ para $i = 0, 1, \ldots, m$ então, para todo $z \in C$ temos $|q(z) - p(z)| \leq \sum_{i=0}^{m} |b_i - a_i|\,|z|^{m-i} < \delta \sum_{i=0}^{m} |z|^{m-i} \leq \delta A = \eta/2$. Isto significa que, para todo $z \in C$, o segmento de reta $[p(z), q(z)]$ não contém a origem. Portanto a homotopia linear $H\colon \mathbb{C} \times [0,1] \to \mathbb{C}$, $H(z,t) = (1-t)p(z) + tq(z)$, cumpre $H(z,t) \neq 0$ para todo $z \in C$ e todo $t \in [0,1]$. Como $H(z,0) = p(z)$ e $H(z,1) = q(z)$, segue-se do Teorema 16 que o número de raízes do polinômio q no disco D (contadas conforme suas multiplicidades) é igual a $n(q;D) = n(p;D) = k$.

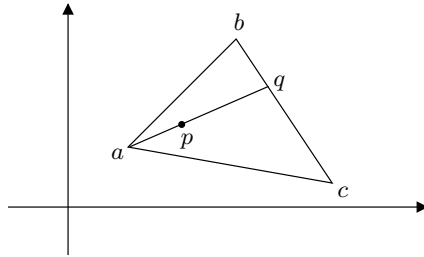
Em particular, se a é uma raiz simples de p, dado qualquer $\varepsilon > 0$ tal que a seja o único zero de p em $B(a;\varepsilon)$, existe $\delta > 0$ tal que se q é um polinômio cujos coeficientes diferem menos de δ dos seus correspondentes em p, então q possui exatamente uma raiz no disco de centro a e raio ε.

APÊNDICE

Transformações afins do plano

Sejam $a, b, c \in \mathbb{R}^2$ três pontos não colineares e T o triângulo que os tem como vértices. Estabeleceremos os seguintes fatos.

Primeiro: *Todo ponto $p \in T$ se exprime, de modo único, como $p = ua + vb + wc$, onde $u, v, w \in [0, 1]$ e $u + v + w = 1$.*

Demonstração: Dado p, seja q o ponto de interseção da reta ap com o lado $[b, c]$. Então $q = (1 - s)b + sc$, com $s \in [0, 1]$ e $p = (1 - t)a + tq$, $t \in [0, 1]$,

logo $p = (1 - t)a + t(1 - s)b + tsc = ua + vb + wc$, onde $u = 1 - t$, $v = t(1 - s)$ e $w = ts$ pertencem a $[0, 1]$ e têm soma igual a 1. Para provar a unicidade, suponha que $ua + vb + wc = \bar{u}a + \bar{v}b + \bar{w}c$, com $u, v, w, \bar{u}, \bar{v}, \bar{w} \in [0, 1]$ e $u + v + w = \bar{u} + \bar{v} + \bar{w} = 1$. Então $w = 1 - u - v$ e $w = 1 - \bar{u} - \bar{v}$, logo

$$ua + vb + (1 - u - v)c = \bar{u}a + \bar{v}b + (1 - \bar{u} - \bar{v})c$$

e daí $u(a - c) + v(b - c) = \bar{u}(a - c) + \bar{v}(b - c)$.
Como $a - c$ e $b - c$ são linearmente independentes, segue-se que $u = \bar{u}$, $v = \bar{v}$ e $w = 1 - u - v = 1 - \bar{u} - \bar{v} = \bar{w}$.

Segundo: *Se $p = ua + vb + wc$, com $u, v, w \in [0, 1]$ e $u + v + w = 1$, então $p \in T$.*

Demonstração: Se $v + w = 0$ então $p = a \in T$. Se $v + w \neq 0$, podemos escrever

$$p = ua + (v + w)\left[\frac{v}{v + w}\, b + \frac{w}{v + w}\, c\right] = ua + (v + w)q.$$

O ponto $q = \dfrac{v}{v+w}\,b + \dfrac{w}{v+w}\,c$ pertence ao lado $[b,c]$, logo $q \in T$. Como $p \in [a,q]$ e o triângulo T é convexo, segue-se que $p \in T$.

Terceiro: *Dados arbitrariamente $A, B, C \in \mathbb{R}^n$, a aplicação $f\colon T \to \mathbb{R}^n$, definida por $f(ua+vb+wc) = uA+vB+wC$, é afim. Se $g\colon T \to \mathbb{R}^n$ é afim e $g(a) = A$, $g(b) = B$, $g(c) = C$, então $g = f$.*

Demonstração: Dados arbitrariamente $p, q \in T$ e $t \in [0,1]$, temos $p = ua+vb+wc$ e $q = \bar{u}a+\bar{v}b+\bar{w}c$, com $u, v, w \in [0,1]$, $u+v+w = 1$, $\bar{u}, \bar{v}, \bar{w} \in [0,1]$ e $\bar{u}+\bar{v}+\bar{w} = 1$. Então:

$$\begin{aligned}
&f((1-t)p + tq)\\
&= f((1-t)ua + (1-t)vb + (1-t)wc + t\bar{u}a + t\bar{v}b + t\bar{w}c)\\
&= f[((1-t)u + t\bar{u})a + ((1-t)v + t\bar{v})b + ((1-t)w + t\bar{w})c]\\
&= ((1-t)u + t\bar{u})A + ((1-t)v + t\bar{v})B + ((1-t)w + t\bar{w})C\\
&= (1-t)(uA + vB + wC) + t(\bar{u}A + \bar{v}B + \bar{w}C)\\
&= (1-t)f(p) + tf(q).
\end{aligned}$$

Isto mostra que f é afim. Evidentemente, $f(a) = A$, $f(b) = B$ e $f(c) = C$. Seja agora $g\colon T \to \mathbb{R}^n$ afim, com $g(a) = A$, $g(b) = B$ e $g(c) = C$. Dado $p = ua + vb + wc \in T$, se for $v + w = 0$ então $p = a$ e daí $g(p) = A = f(p)$. Se, porém, $t = v + w \neq 0$, escrevamos $q = (v/t)b + (w/t)c$. O ponto q pertence a T, pois $v/t + w/t = 1$. Sendo g afim, temos $g(q) = (v/t)B + (w/t)C = f(q)$. Portanto $g(p) = g((1-t)a + tq) = (1-t)A + t \cdot g(q) = (1-t)f(a) + tf(q) = f((1-t)a + tq) = f(p)$. Assim, $g = f$, como queríamos demonstrar.

Exercícios

§2.

2.1. Seja $\alpha\colon [a,b] \to \mathbb{R}$ constante nos intervalos $[a,c)$ e $(c,b]$, onde $a < c < b$. Mostre que se $f\colon [a,b] \to \mathbb{R}$ é contínua no ponto c então $\int_a^b f d\alpha = f(c)\cdot[\alpha(c^+) - \alpha(c^-)]$. Por outro lado, se f e α são descontínuas no ponto c então não existe a integral $\int_a^b f d\alpha$. Em particular, dada $f\colon [a,b] \to \mathbb{R}$, se existe a integral $\int_a^b f d\alpha$ para toda função $\alpha\colon [a,b] \to \mathbb{R}$ de variação limitada, então f é contínua.

2.2. Seja $\alpha\colon [a,b] \to \mathbb{R}$ uma função limitada. Se existe $\int_a^b f d\alpha$ para toda função contínua $f\colon [a,b] \to \mathbb{R}$, então α é de variação limitada.

2.3. Sejam $f, \alpha\colon [a,b] \to \mathbb{R}$ limitadas, tais que a integral $\int_a^b f d\alpha$ existe. Então, para todo $c \in [a,b]$, existem as integrais $\int_a^c f d\alpha$, $\int_c^b f d\alpha$ e vale $\int_a^b f d\alpha = \int_a^c f d\alpha +$

$\int_c^b f d\alpha$. Considere $f, \alpha\colon [-1,1] \to \mathbb{R}$ definidas por $f(x) = 0$ se $-1 \leq x < 0$, $f(x) = 1$ se $0 \leq x \leq 1$, $\alpha(x) = 0$ se $-1 \leq x \leq 0$ e $\alpha(x) = 1$ se $0 < x \leq 1$. Mostre que existem $\int_{-1}^0 f d\alpha$ e $\int_0^1 f d\alpha$ mas não existe $\int_{-1}^1 f d\alpha$.

2.4. Sejam $f, \alpha\colon [a,b] \to \mathbb{R}$ limitadas. Existe $\int_a^b f d\alpha$ se, e somente se, existe $\int_a^b \alpha df$. No caso afirmativo, tem-se $\int_a^b f d\alpha + \int_a^b \alpha df = f(b)\alpha(b) - f(a)\alpha(a)$.

2.5. Seja $\alpha\colon (0,+\infty) \to (0,+\infty)$ contínua crescente, com $\alpha(1) = 1$. Prove que $\int_1^x d\alpha/\alpha = \log \alpha(x)$. [Não suponha α diferenciável!]

2.6. Sejam $f\colon U \times [a,b] \to \mathbb{R}$ como na Regra de Leibniz (Cap. III, §6) e $\alpha\colon [a,b] \to \mathbb{R}$ de variação total limitada. Prove que se

$$F(x) = \int_a^b f(x,t)d\alpha(t) \text{ então } \frac{\partial F}{\partial x_i}(x) = \int_a^b \frac{\partial f}{\partial x_i}(x,t)d\alpha(t),$$

para todo $x \in U$.

2.7. Sejam f contínua e α de variação limitada no intervalo $[a,b]$. Defina $F\colon [a,b] \to \mathbb{R}$ pondo $F(x) = \int_a^x f d\alpha$. Prove que F é de variação limitada e que a variação total de F não excede $M \cdot V$, onde $M = \sup\{|f(x)|; x \in [a,b]\}$ e V é a variação total de α em $[a,b]$.

2.8. Dada uma seqüência de funções limitadas $f_n\colon [a,b] \to \mathbb{R}$, suponha que exista, para cada n, a integral $\int_a^b f_n(t)d\alpha$, onde $\alpha\colon [a,b] \to \mathbb{R}$ tem variação limitada. Suponha ainda que $f_n \to f$ uniformemente em $[a,b]$. Prove que existe $\int_a^b f d\alpha$ e vale $\int_a^b f d\alpha = \lim \int_a^b f_n d\alpha$. Dê exemplo em que $\alpha_n \to \alpha$ uniformemente mas não se tem $\int_a^b f d\alpha_n \to \int_a^b f d\alpha$.

§3.

3.1. Se o caminho $\lambda\colon [a,b] \to \mathbb{R}^2$ é o gráfico de uma função $f\colon [a,b] \to \mathbb{R}$ de classe C^1, isto é, $\lambda(x) = (x, f(x))$, e $\omega = Adx + Bdy$ é uma 1-forma contínua definida num conjunto que contém a imagem de λ então

$$\int_\lambda \omega = \int_a^b [A(x,f(x)) + B(x,f(x)), f'(x)]dx.$$

3.2. Se $\lambda\colon [a,b] \to U \subset \mathbb{R}^2$ é o caminho retilíneo horizontal, que liga o ponto (a,c) ao ponto (b,c) e $\omega = fdx + gdy$ é uma 1-forma em U, a integral $\int_\lambda \omega$ existe se, e somente se, a função $x \mapsto f(x,c)$ é integrável no intervalo $[a,b]$. No caso afirmativo, tem-se $\int_\lambda \omega = \int_a^b f(x,c)dx$.

3.3. Seja $U \subset \mathbb{R}^2$ um aberto contendo o círculo unitário S^1. Para cada $k \in \mathbb{Z}$, considere o caminho $\lambda_k\colon [0,2\pi] \to U$, dado por $\lambda_k(t) = (\cos kt, \operatorname{sen} kt)$. Se $\omega\colon U \to (\mathbb{R}^2)^*$ é qualquer 1-forma contínua, prove que $\int_{\lambda_k} \omega = k\int_{\lambda_1} \omega$.

3.4. Para cada $t \in T \subset \mathbb{R}^m$, seja $\omega_t = \sum_{i=1}^n a_i(t,x)dx_i$ uma forma diferencial no aberto $U \subset \mathbb{R}^n$, tal que os coeficientes $a_i\colon T \times U \to \mathbb{R}$ são funções contínuas. Para cada caminho retificável $\lambda\colon [a,b] \to U$, prove que a função $\varphi\colon T \to \mathbb{R}$, definida por $\varphi(t) = \int_\lambda \omega_t$, é uma função contínua.

3.5. Calcule $\int_\lambda f ds$ onde $f\colon \mathbb{R}^2 \to \mathbb{R}$ é dada por $f(x,y) = \sqrt{x^2+y^2}$ e $\lambda\colon [0,2\pi] \to \mathbb{R}^2$ é o caminho $\lambda(t) = (a\cos t, b \operatorname{sen} t)$, que descreve uma elipse.

3.6. Seja $\lambda\colon I \to U$ um caminho contínuo retificável. Se uma seqüência de 1-formas ω_n, contínuas no aberto $U \subset \mathbb{R}^m$, converge uniformemente em $\lambda(I)$ para uma forma ω (isto é, cada coeficiente de ω_n converge uniformemente para o coeficiente correspondente em ω), então $\lim\limits_{n\to\infty} \int_\lambda \omega_n = \int_\lambda \omega$.

3.7. Sejam $\lambda\colon [a,b] \to \mathbb{R}^2 - \{0\}$ um caminho fechado seccionalmente C^1 e $\alpha \in \mathbb{R}$ uma constante. Defina $\mu\colon [a,b] \to \mathbb{R}^2 - \{0\}$ pondo $\mu(t) = e^{i\alpha} \cdot \lambda(t) = (x(t)\cos\alpha - y(t)\operatorname{sen}\alpha, x(t)\operatorname{sen}\alpha + y(t)\cos\alpha)$, onde $\lambda(t) = (x(t), y(t))$. Os caminhos λ e μ dão o mesmo número de voltas em torno da origem.

3.8. Sejam $\lambda, \mu\colon [a,b] \to \mathbb{R}^2 - \{0\}$ caminhos fechados de classe C^1. Se $\lambda(t)$ e $\mu(t)$ são linearmente independentes para todo $t \in [a,b]$ então λ e μ dão o mesmo número de voltas em torno da origem.

3.9. Dado o caminho fechado $\lambda\colon [a,b] \to \mathbb{R}^2 - \{0\}$, de classe C^1, indique com $A(t)$ a "área orientada" do paralelogramo determinado pelos vetores $\lambda(t)$ e $\lambda'(t)$, para cada $t \in [a,b]$. Mostre que o número de voltas que λ percorre em torno da origem é

$$n(\lambda; 0) = \frac{1}{2\pi} \int_a^b \frac{A(t)}{|\lambda(t)|^2}\, dt.$$

[A "área orientada" de um paralelogramo cujos lados são dados pelos vetores $v, w \in \mathbb{R}^2$ (nesta ordem) é positiva quando o sentido de rotação $v \mapsto w$ pelo menor ângulo é anti-horário, negativa no caso contrário e nula quando v e w são linearmente dependentes.]

3.10. Se o caminho fechado $\lambda\colon [a,b] \to \mathbb{R}^2 - \{0\}$, de classe C^1, é tal que $\lambda(t)$ e $\lambda'(t)$ são linearmente independentes para todo t (ou, mais geralmente, quando o determinante 2×2 cujas colunas são $\lambda(t)$ e $\lambda'(t)$ não muda de sinal quando t percorre $[a,b]$) então o número de voltas que λ dá em torno da origem é $\neq 0$.

3.11. (*Teorema de Green para um retângulo.*) Parametrize o bordo do retângulo $R = [A,B] \times [C,D]$ como um caminho ∂R, seccionalmente de classe C^1, percorrido no sentido anti-horário. [Se $\partial R(t) = (x(t), y(t))$, isto significa que $x' > 0$ quando $y = C$, $x' < 0$ quando $y = D$, $y' > 0$ quando $x = B$ e $y' < 0$ se $x = A$.] Seja $\omega = adx + bdy$ uma 1-forma de classe C^1 num aberto do plano que contém o retângulo R. Mostre que

$$\iint_R \left(\frac{\partial b}{\partial x} - \frac{\partial a}{\partial y}\right) dxdy = \int_{\partial R} adx + bdy,$$

onde $\iint_R$ significa a integral repetida $\int_C^D \int_A^B$.

3.12. (*Teorema de Green num disco.*) Seja $\omega = adx + bdy$ uma 1-forma de classe C^1 num aberto $U \subset \mathbb{R}^2$, que contém o disco fechado D, de centro $p = (x_0, y_0)$ e raio r. A fronteira ∂D será identificada ao caminho $t \mapsto (x_0 + r\cos t, y_0 + r \operatorname{sen} t)$. Mostre que

$$\int_{\partial D} adx + bdy = \iint_D \left(\frac{\partial b}{\partial x} - \frac{\partial a}{\partial y}\right) dxdy,$$

onde o segundo membro é a integral repetida

$$\int_{x_0-r}^{x_0+r} dx \int_{y_0-\sqrt{r^2-(x-x_0)^2}}^{y_0+\sqrt{r^2-(x-x_0)^2}} \left(\frac{\partial b}{\partial x} - \frac{\partial a}{\partial y}\right) dy.$$

§4.

4.1. Sejam $\varphi, \psi\colon [a,b] \to \mathbb{R}$ de classe C^1, com $\varphi(a) = \psi(a)$, $\varphi(b) = \psi(b)$, $\varphi(x) \leq \psi(x)$ para todo $x \in [a,b]$, φ convexa e ψ côncava. Seja λ o caminho fechado obtido por justaposição de $t \mapsto (t, \varphi(t))$, gráfico de φ, com o oposto de $t \mapsto (t, \psi(t))$, gráfico de ψ. Se $f dx + g dy$ é uma forma de classe C^1, definida num aberto convexo $U \subset \mathbb{R}^2$ que contém a imagem de λ, prove que

$$\int_\lambda f dx + g dy = \int_a^b dx \int_{\varphi(x)}^{\psi(x)} \left(\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y}\right) dy.$$

§6.

6.1. Seja ω uma 1-forma de classe C^1 na bola aberta $B \subset \mathbb{R}^m$, tal que a integral $\int_\lambda \omega$ assume um valor constante para qualquer caminho fechado λ, seccionalmente C^1, contido em B. Então ω é fechada.

6.2. Seja ω uma 1-forma de classe C^1 num aberto $U \subset \mathbb{R}^m$. Suponha que, para todo caminho fechado λ, seccionalmente C^1 em U, a integral $\int_\lambda \omega$ é um número racional. Prove que a forma ω é fechada.

6.3. Sejam $U \subset \mathbb{R}^m$ aberto e $A\colon U \to \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$ contínua. Para todo caminho $\lambda\colon [a,b] \to U$ de classe C^1, defina a integral de A ao longo do caminho λ como o vetor de $\mathbb{R}^n$ dado por $\int_\lambda A = \int_a^b A(\lambda(t)) \cdot \lambda'(t) dt$. Estenda esta definição para caminhos seccionalmente C^1. Seja $(a_{ij}(x))$ a matriz de $A(x)$ nas bases canônicas de $\mathbb{R}^m$ e $\mathbb{R}^n$. Prove que $\int_\lambda A = 0$ para todo caminho fechado λ em U se, e somente se, existem n funções $f_1, \ldots, f_n\colon U \to \mathbb{R}$, de classe C^1, tais que $a_{ij} = \partial f_i / \partial x_j$.

6.4. Sejam ω uma 1-forma fechada, diferente de zero em todos os pontos do aberto $U \subset \mathbb{R}^m$, e $f\colon U \to \mathbb{R}$ uma função de classe C^1 sem pontos críticos. As seguintes afirmações são equivalentes: (a) a forma $f \cdot \omega$ é fechada; (b) df é um múltiplo de ω; (c) para todo $x \in U$, e todo $v \in \mathbb{R}^m$, $\omega(x) \cdot v = 0 \Leftrightarrow \langle \operatorname{grad} f(x), v \rangle = 0$.

§7.

7.1. Uma homotopia $H\colon [a,b] \times I \to V$ (tomando valores no aberto $V \subset \mathbb{R}^n$) diz-se *de classe* C^r quando $H(s,t) = (x_1(s,t), \ldots, x_n(s,t))$, onde $x_i\colon U \to \mathbb{R}$ são funções de classe C^r num aberto U do plano, contendo o retângulo $[a,b] \times I$. Suponha que exista uma homotopia H de classe C^2 (com extremidades fixas, naturalmene) entre dois caminhos $\lambda, \mu\colon [a,b] \to V$. Suponha ainda que ω seja uma 1-forma fechada definida em V. Use o Teorema de Green, aplicado à 1-forma $H^*\omega$ sobre o retângulo $[a,b] \times I$, para mostrar que $\int_\lambda \omega = \int_\mu \omega$. Isto fornece uma nova prova para o Teorema 12 no caso de uma homotopia de classe C^2.

§8.

8.1. Dado $p = (a, b) \in \mathbb{R}^2$, mostre que a 1-forma $\omega_p \colon \mathbb{R}^2 - \{p\} \to (\mathbb{R}^2)^*$, definida por $\omega_p(x, y) = \dfrac{(x-a)dy - (y-b)dx}{(x-a)^2 + (y-b)^2}$, é fechada. Prove que esta forma é exata num aberto $U \subset \mathbb{R}^2 - \{p\}$ se, e somente se, existe uma função contínua (necessariamente C^∞) $\theta_p \colon U \to \mathbb{R}$ tal que $\cos\theta_p(z) = \dfrac{x-a}{|z-p|}$ e $\operatorname{sen}\theta_p(z) = \dfrac{y-b}{|z-p|}$ para todo ponto $z = (x, y) \in U$. Uma função θ_p com estas propriedades chama-se uma *função-ângulo de vértice p.*

8.2. Para cada ponto $p = (a, b) \in \mathbb{R}^2$, considere a 1-forma ω_p em $R^2 - \{p\}$, dada por $\omega_p(x, y) = \dfrac{(x-a)dy - (y-b)dx}{(x-a)^2 + (y-b)^2}$ e, para cada caminho contínuo retificável $\lambda \colon I \to \mathbb{R}^2 - \{p\}$, seja $\lambda_p \colon I \to \mathbb{R}^2 - \{0\}$, com $\lambda_p(t) = \lambda(t) - p$. Mostre que $\int_\lambda \omega_p = \int_{\lambda_p} \delta\theta$. Conclua que se λ é fechado e os pontos p, q podem ser ligados por um caminho contínuo em $\mathbb{R}^2 - \lambda(I)$ então $\int_\lambda \omega_p = \int_\lambda \omega_q$.

8.3. Sejam $p = (a, b)$ e $q = (c, d)$ pontos na mesma componente conexa por caminhos de $\mathbb{R}^2 - U$, $U \subset \mathbb{R}^2$ aberto. A forma

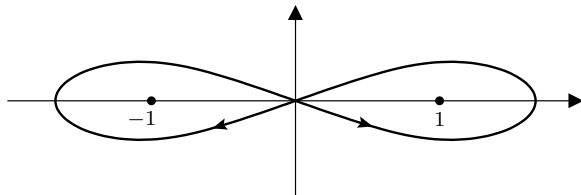
$$\omega_p = \frac{(x-a)dy - (y-b)dx}{(x-a)^2 + (y-b)^2}$$

é exata em U se, e somente se, é também exata em U a forma

$$\omega_q = \frac{(x-c)dy - (y-d)dx}{(x-c)^2 + (y-d)^2}.$$

Mais precisamente, mostre que existe uma função $f \colon U \to \mathbb{R}$, de classe C^∞, tal que $\omega_p - \omega_q = df$. [*Sugestão*: mostre que $\int_\lambda \omega_p = \int_\lambda \omega_q$ para todo caminho fechado λ, seccionalmente C^1 em U.]

8.4. Se o aberto $U \subset \mathbb{R}^2$ é limitado, mostre que, para todo $p \in \mathbb{R}^2$ de norma suficientemente grande, U está contido no complementar de uma semi-reta de origem p, logo existe uma função-ângulo contínua $\theta_p \colon U \to \mathbb{R}$, de vértice p. Conclua que a forma ω_p é exata em U. Em particular, se $U \subset \mathbb{R}^2$ é limitado e $\mathbb{R}^2 - U$ é conexo por caminhos então, para todo $p \in \mathbb{R}^2 - U$, a forma ω_p é exata em U. Isto nos dá uma função-ângulo $\theta \colon U \to \mathbb{R}$ em todo aberto limitado $U \subset \mathbb{R}^2 - \{0\}$ tal que $\mathbb{R}^2 - U$ é conexo.

8.5. Calcule $\int_\lambda dz/(z^2-1)^2$ onde λ é o caminho em $\mathbb{R}^2 - \{1, -1\}$ esboçado a seguir.

§9.

9.1. Seja $X = (a, b, c)\colon U \to \mathbb{R}^3$ um campo de vetores de classe C^∞ em $\mathbb{R}^3$, com

$$\frac{\partial a}{\partial y} = \frac{\partial b}{\partial x}, \ \frac{\partial b}{\partial z} = \frac{\partial c}{\partial y} \quad \text{e} \quad \frac{\partial a}{\partial z} = \frac{\partial c}{\partial x}$$

em todos os pontos do aberto simplesmente conexo $U \subset R^3$. Mostre que existe uma função $f\colon U \to \mathbb{R}$ de classe C^∞ com a seguinte propriedade: $df(p) \cdot v = 0$ se, e somente se, $\langle v, X(p)\rangle = 0$.

9.2. Seja $\delta\theta$ a forma elemento de ângulo. Para todo $z \in \mathbb{R}^2 - \{0\}$, seja $\hat{z}$ o vetor obtido de z por rotação de $90°$ no sentido anti-horário. Tem-se $\delta\theta(z) \cdot v = \langle v, \hat{z}\rangle/|z|^2$.

9.3. Seja $\lambda\colon [a, b] \to \mathbb{R}^2 - \{0\}$ um caminho regular fechado de classe C^1, tal que $\langle \widehat{\lambda(t)}, \lambda'(t)\rangle \geq 0$ para todo $t \in [a, b]$ (na notação do exercício anterior). Seja ainda ω uma 1-forma fechada em $\mathbb{R}^2 - \{0\}$ tal que, para todo $t \in [a, b]$, $\lambda'(t)$ pertence ao núcleo de $\omega(\lambda(t))$. Prove que ω é exata.

9.4. Uma 1-forma ω, fechada em $\mathbb{R}^2 - \{0\}$ e limitada numa vizinhança da origem (isto é, seus coeficientes são limitados numa bola de centro 0) é exata. Mostre que o resultado ainda vale se supusermos apenas que $\lim\limits_{z \to 0} |z| \cdot \omega(z) = 0$. No primeiro caso, mostre que existe sempre $f\colon \mathbb{R}^2 \to \mathbb{R}$ contínua, com $df = \omega$ em $\mathbb{R}^2 - \{0\}$.

9.5. Seja ω uma 1-forma fechada num aberto $U \subset \mathbb{R}^2$, complementar de uma bola fechada centrada na origem. Se $\lim\limits_{|z| \to \infty} |z|\omega(z) = 0$ então ω é exata. Conclua que a 1-forma fechada ω é exata desde que se anule para todo z de módulo suficientemente grande. Em particular, se ω tem suporte compacto então ω é exata. [O *suporte* de uma forma ω é o fecho do conjunto dos pontos z tais que $\omega(z) \neq 0$.]

9.6. Mostre que a forma diferencial definida em $\mathbb{R}^2 - \{0\})$

$$\omega = \frac{e^x}{x^2 + y^2}\,[(x \cos y + y \operatorname{sen} y)dy + (x \operatorname{sen} y - y \cos y)dx]$$

se escreve como $\omega = e^x \cdot \cos y \cdot \delta\theta + e^x \cdot \operatorname{sen} y \cdot d(\log r)$, onde $r(x, y) = \sqrt{x^2 + y^2}$. Verifique que ω é fechada e que $\alpha = \omega - \delta\theta$ cumpre a condição $\lim\limits_{z \to 0} |z|\alpha(z) = 0$. Conclua que α é exata em $\mathbb{R}^2 - \{0\}$. A partir daí, calcule $\int_\lambda \omega$ para qualquer caminho fechado λ de classe C^1 em $\mathbb{R}^2 - \{0\}$.

9.7. Prove que toda função harmônica em $\mathbb{R}^2 - \{0\}$ é igual à parte real de uma função holomorfa mais um múltiplo constante da função $\log \sqrt{x^2 + y^2}$.

9.8. Seja $U = \mathbb{R}^m - F$, onde $m > 2$ e F é um conjunto finito. Toda 1-forma fechada em U é exata.

9.9. Seja $U = \{(x, y, z) \in \mathbb{R}^3, x^2 + y^2 \neq 0\}$ o complementar do eixo z. Mostre que toda forma fechada em U se escreve como $c \cdot \alpha + df$, onde $c \in \mathbb{R}$, $f\colon U \to \mathbb{R}$ é uma função de classe C^2 e $\alpha(x, y, z) = \dfrac{xdy - ydx}{x^2 + y^2} \cdot$ Em particular, $H^1(U)$ tem dimensão 1.

9.10. Dado um retângulo $R = [a, b] \times [c, d] \subset \mathbb{R}^2$, indique com $\partial R = (x(t), y(t))$ o caminho poligonal fechado que descreve o bordo de R no sentido anti-horário, isto é, $x' > 0$ se $y = c$, $x' < 0$ se $y = d$, $y' > 0$ se $x = b$ e $y' < 0$ se $x = a$. Fixe um retângulo $R_0 \subset \mathbb{R}^2$ e considere uma 1-forma fechada ω, definida em $U = \mathbb{R}^2 - \partial R_0$. Se, para algum retângulo R tal que $\partial R \subset U$, com o mesmo centro de R_0, tem-se $\int_{\partial R} \omega = 0$, prove que existe uma função $f\colon U \to \mathbb{R}$, de classe C^2, tal que $df = \omega$.

9.11. Sejam $U, V \subset \mathbb{R}^m$ abertos simplesmente conexos. Se $U \cap V$ é conexo, então a reunião $U \cup V$ é simplesmente conexa. Conclua daí uma nova demonstração de que, para $m > 1$, o conjunto $\mathbb{R}^{m+1} - \{0\}$ é simplesmente conexo, logo a esfera S^m também o é.

9.12. O *suporte* de uma aplicação $f\colon X \to \mathbb{R}^n$, definida num conjunto $X \subset \mathbb{R}^m$, é o conjunto $\text{supp}\cdot f$, formado pelos pontos $x \in X$ tais que $x = \lim x_k$ com $x_k \in X$ e $f(x_k) \neq 0$ para todo $k \in \mathbb{N}$. Seja $f\colon \mathbb{R} \to \mathbb{R}$ uma função contínua com suporte compacto. Prove que, a fim de existir uma função $g\colon \mathbb{R} \to \mathbb{R}$, de classe C^1, com suporte compacto, tal que $g' = f$, é necessário e suficiente que $\int_{-\infty}^{+\infty} f(x)dx = 0$. [A integral infinita significa, neste caso, apenas $\int_a^b f(x)dx$, onde $[a, b]$ é qualquer intervalo compacto contendo o suporte de f.]

9.13. Sejam α e β 1-formas fechadas com suporte compacto, definidas no aberto $U \subset \mathbb{R}^2$. Escreva $\alpha \sim \beta$ para indicar que $\alpha - \beta = df$, onde $f\colon U \to \mathbb{R}$ é uma função de classe C^2 *com suporte compacto*, definida em U. Mostre que a relação $f \sim g$ é uma eqüivalência no conjunto das 1-formas fechadas com suporte compacto definidas em U e que as classes de equivalência daí resultantes formam, de modo natural, um espaço vetorial $H_c^1(U)$, chamado o "grupo de cohomologia real de dimensão 1 com suporte compacto"do aberto U. Mostre que $dim.H_c^1(\mathbb{R}^2) = 0$. Mostre também que se $U \subset \mathbb{R}^2$ é um aberto limitado tal que $\mathbb{R}^2 - U$ tem $n+1$ componentes conexas então $\dim .H_c^1(U) = n$. Finalmente, se U é um aberto tal que $\mathbb{R}^2 - U$ é limitado e tem $n+1$ componentes conexas então $\dim .H_c^1(U) = n$. [*Sugestão*: Toda 1-forma fechada de classe C^1, com suporte compacto, definida em U é a restrição de uma forma fechada de classe C^1 definida em $\mathbb{R}^2$, igual a zero em $\mathbb{R}^2 - U$.]

9.14. Seja $F = (a, b)\colon U \to \mathbb{R}^2$ um campo de vetores de classe C^1 no aberto $U \subset \mathbb{R}^2$. A *divergência* e o *rotacional* de F são as funções reais

$$\text{div}\,.F = \frac{\partial a}{\partial x} + \frac{\partial b}{\partial y}, \quad \text{rot}.F = \frac{\partial b}{\partial x} - \frac{\partial a}{\partial y},$$

definidas em U. Uma função $f\colon U \to \mathbb{R}$ é a divergência de um campo de vetores em U se, e somente se, é o rotacional de um outro campo. Se U é convexo, prove que toda função contínua $f\colon U \to \mathbb{R}$ é a divergência de algum campo de vetores $F\colon U \to \mathbb{R}^2$. (Se uma das derivadas parciais de f existe e é contínua em U, o campo F pode ser tomado de classe C^1.)

9.15. Seja $f\colon U \to \mathbb{R}$ uma função contínua com suporte compacto, definida num aberto $U \subset \mathbb{R}^2$. Observe que f é a restrição de uma função contínua definida em todo o plano e conclua que f é a divergência de algum campo de vetores $F\colon U \to \mathbb{R}^2$. A fim de que exista um campo $F\colon U \to \mathbb{R}^2$ de classe C^1 e suporte compacto, tal que $f = \text{div}\,.F$, é necessário que $\iint f(x, y)dxdy = 0$. [Esta

integral dupla significa $\int_\alpha^\beta dx \int_\gamma^\delta f(x,y)dy$, onde f é suposta estendida a todo o plano, igual a zero fora de U, e $[\alpha,\beta]\times[\gamma,\delta]$ é qualquer retângulo contendo o suporte de f.]

9.16. Seja $f\colon \mathbb{R}^2\to\mathbb{R}$ uma função de classe C^1 com suporte compacto. A condição $\iint f(x,y)dxdy = 0$ é suficiente para que seja $f = \text{div}\,.F$ onde $F\colon \mathbb{R}^2\to\mathbb{R}$ é um campo de vetores de classe C^1 com suporte compacto. [Seja $Q=\{(x,y)\in\mathbb{R}^2; |x|\le k, |y|\le k\}$ um quadrado contendo o suporte de f. Pondo $a(x,y) = \int_{-k}^x f(s,y)ds$, veja que $a\colon \mathbb{R}^2\to\mathbb{R}$ é de classe C^1 e $\partial a/\partial x = f$. A 1-forma $\omega = a dy$ é fechada em $U=\mathbb{R}^2 - Q$ e $\int_{\partial R}\omega = 0$ para todo retângulo $R\supset Q$. Conclua que existe $g\colon U\to\mathbb{R}$ de classe C^2 tal que $\omega = dg$ em U. A função g não está definida em todo o plano, mas contorne esta dificuldade tomando $\varphi\colon \mathbb{R}\to\mathbb{R}$ de classe C^2 tal que $\varphi(t)=1$ se $|t|\le 2k$ e $\varphi(t)=0$ se $|t|\ge 3k$. Defina $h\colon \mathbb{R}^2\to\mathbb{R}$ pondo $h(x,y)=0$ se $(x,y)\in Q$ e $h(x,y) = [1-\varphi(x)\varphi(y)]\cdot g(x,y)$ se $(x,y)\in U$. Agora vale $dh=\omega$ no aberto $V = R^2 - Q_1$, onde $Q_1 = \{(x,y)\in\mathbb{R}^2; |x|\le 3k, |y|\le 3k\}$. Finalmente, tome $A = a - \partial h/\partial y$, $B = \partial h/\partial x$. Mostre que $\partial A/\partial x + \partial B/\partial y = f$ em $\mathbb{R}^2$ e $A = B = 0$ em V, logo $F=(A,B)$ tem suporte compacto.]

9.17. Seja $U\subset\mathbb{R}^2$ aberto. Dadas duas funções $f,g\colon U\to\mathbb{R}$, de classe C^1, com suporte compacto, escreva $f\sim g$ para indicar que $f-g=\text{div}\,.F$ onde $F\colon U\to\mathbb{R}^2$ é algum campo vetorial de classe C^1 com suporte compacto. Mostre que a relação $f\sim g$ é uma equivalência no conjunto das funções de classe C^1 com suporte compacto em U e que as classes de equivalência daí resultantes constituem, de modo natural, um espaço vetorial $H_c^2(U)$, chamado o "grupo de cohomologia real de dimensão 2 com suporte compacto" do aberto U. Mostre que a aplicação $f\mapsto \iint f\,dxdy$ é constante em cada classe, logo induz um homomorfismo (funcional linear) de $H_c^2(U)$ sobre $\mathbb{R}$. Mostre que se $U=\mathbb{R}^2$ então esse homomorfismo é injetivo, logo $\dim H_c^2(\mathbb{R}^2) = 1$. [Observação: Este resultado vale, em geral, para todo aberto conexo $U\subset\mathbb{R}^2$.]

§10.

10.1. Seja $F\colon U\to\mathbb{R}^2$ de classe C^2 no aberto $U\subset\mathbb{R}^2$, contendo o disco D. Se $0\notin F(\partial D)$, prove que $n(F;D) = n\left(\dfrac{F}{|F|};D\right)$.

10.2. Sejam $F,G\colon U\to\mathbb{R}^2$ aplicações de classe C^2 no aberto $U\subset\mathbb{R}^2$ e $D\subset U$ um disco fechado. Se $0<|G(z)|<|F(z)|$ para qualquer $z\in\partial D$ então $n(F+G;D) = n(F;D)$.

10.3. Seja $F\colon U\to\mathbb{R}^2$ uma aplicação de classe C^2 no aberto $U\subset\mathbb{R}^2$, o qual contém o disco fechado D de centro 0. Se $F(-z) = -F(z)\ne 0$ para todo $z\in\partial D$ então $n(F;D)$ é um número ímpar. Em particular, deve existir $z\in D$ tal que $F(z)=0$.

10.4. O *polinômio característico* de uma matriz $A\in\mathbb{R}^{m^2}$ é o polinômio, de grau m, $p(\lambda) = \det.(A-\lambda I)$, onde I é a matriz identidade $m\times m$. O *espectro* de A é o conjunto $\text{Esp}(A)$ das raízes (reais ou complexas) do polinômio característico de A. Uma matriz A chama-se *hiperbólica* quando $\text{Esp}(A)\cap S^1=\varnothing$, isto é, quando nenhuma das raízes do seu polinômio característico é um número complexo de

módulo 1. Prove que o conjunto das matrizes hiperbólicas é aberto em $\mathbb{R}^{m^2}$. Observe ainda que, se ε é qualquer número real, o espectro da matriz $A - \varepsilon I$ é obtido do espectro de A por uma translação horizontal de comprimento ε. Conclua daí que o conjunto das matrizes hiperbólicas, além de aberto, é denso em $\mathbb{R}^{m^2}$.

10.5. Mostre que o sistema de equações

$$\begin{aligned} 3x^2 - 2xy + y^2 &= 2 \\ 4x^2 + 2xy + 2y^2 &= 3 \end{aligned}$$

possui uma solução no interior do disco unitário em $\mathbb{R}^2$.

10.6. Se $|a_k| > \sum_{j \neq k} |a_j|$ para algum $k \in [0, n]$ então o polinômio $p(z) = a_0 + a_1 z + \cdots + a_n z^n$ possui exatamente k raízes (contadas de acordo com suas multiplicidades) no disco unitário $|z| < 1$.

10.7. Se $F\colon U \to \mathbb{C}$ é holomorfa e $D \subset U$ é um disco em cujo bordo C, F não se anula, então

$$n(F; D) = \frac{1}{2\pi i} \int_c \frac{F'(z)}{F(z)}\, dz.$$

Capítulo V

Aplicações Diferenciáveis

1 Diferenciabilidade de uma aplicação

Uma aplicação f é diferenciável no ponto a quando, para pequenos valores de v, o acréscimo $f(a+v) - f(a)$ é, aproximadamente, uma função linear de v. Mais precisamente:

A aplicação $f \colon U \to \mathbb{R}^n$, definida no aberto $U \subset \mathbb{R}^m$, diz-se *diferenciável* no ponto $a \in U$ quando existe uma aplicação linear $T \colon \mathbb{R}^m \to \mathbb{R}^n$ tal que

$$f(a+v) - f(a) = T \cdot v + r(v), \text{ onde } \lim_{v \to 0} \frac{r(v)}{|v|} = 0.$$

Aqui, supõe-se tacitamente que $a + v \in U$, para que $f(a+v)$ tenha sentido. Como U é aberto, existe $\delta > 0$ tal que $|v| < \delta \Rightarrow a + v \in U$. A igualdade acima é a definição do "resto" $r(v)$. Uma vez dada T, a diferenciabilidade de f no ponto a tem sua essência na afirmação de que $r(v)$ é infinitésimo em relação a v, o que se exprime com $\lim_{v \to 0} r(v)/|v| = 0$. Ou, em termos explícitos: para todo $\varepsilon > 0$ existe $\delta > 0$ tal que $0 < |v| < \delta \Rightarrow |r(v)| < \varepsilon |v|$.

Em alguns casos, para evitar as exceções causadas pelo denominador zero, é conveniente pôr o resto sob a forma $r(v) = \rho(v) \cdot |v|$, onde a aplicação ρ é definida, para todo v tal que $a+v \in U$, por $\rho(v) = r(v)/|v|$, se $v \neq 0$, e $\rho(0) = 0$. Então a diferenciabilidade de f no ponto a se exprime como

$$f(a+v) - f(a) = T \cdot v + \rho(v) \cdot |v|, \text{ onde } \lim_{v \to 0} \rho(v) = 0,$$

de modo que ρ é contínua no ponto 0.

Como a validez da condição $\lim_{v\to 0} r(v)/|v| = 0$ (ou $\lim_{v\to 0} \rho(v) = 0$) independe das normas tomadas em $\mathbb{R}^m$ e $\mathbb{R}^n$, vemos que o fato de uma aplicação ser ou não diferenciável num determinado ponto também não depende das normas.

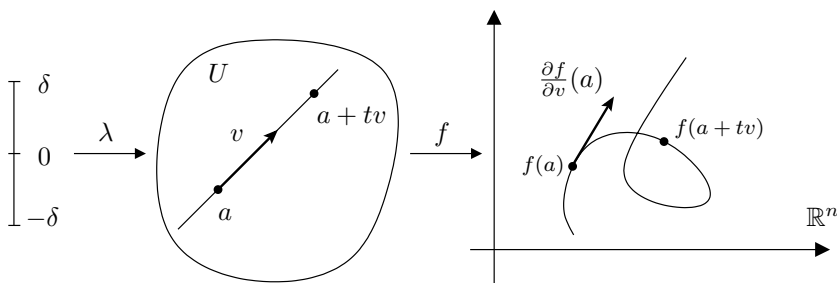
Toda aplicação diferenciável num ponto é, evidentemente, contínua nesse ponto.

Vejamos agora como interpretar a transformação linear $T\colon \mathbb{R}^m \to \mathbb{R}^n$ que ocorre na definição acima.

Seja $f\colon U \to \mathbb{R}^n$ definida num aberto $U \subset \mathbb{R}^m$. A *derivada direcional de* f num ponto $a \in U$, relativamente a um vetor $v \in \mathbb{R}^m$, é, por definição,

$$\frac{\partial f}{\partial v}(a) = \lim_{t\to 0} \frac{f(a+tv) - f(a)}{t} \in \mathbb{R}^n,$$

quando tal limite existe.

Podemos visualizar $\dfrac{\partial f}{\partial v}(a)$ do seguinte modo: seja $\delta > 0$ tal que o segmento de reta aberto $(a-\delta v, a+\delta v)$ esteja contido em U. O caminho retilíneo $\lambda\colon (-\delta, \delta) \to U$, dado por $\lambda(t) = a + tv$, é transformado por f no caminho $f \circ \lambda\colon t \mapsto f(a+tv)$, no espaço $\mathbb{R}^n$. A derivada direcional $\dfrac{\partial f}{\partial v}(a)$ é o vetor-velocidade de $f \circ \lambda$ no instante $t = 0$.

Se $f = (f_1, \ldots, f_n)$ então $\dfrac{\partial f}{\partial v}(a) = \left(\dfrac{\partial f_1}{\partial v}(a), \ldots, \dfrac{\partial f_n}{\partial v}(a)\right)$. Quando $v = e_j$ é o j-ésimo vetor da base canônica de $\mathbb{R}^m$, escreve-se $\dfrac{\partial f}{\partial x_j}(a)$ em vez de $\dfrac{\partial f}{\partial e_j}(a)$. Assim, $\dfrac{\partial f}{\partial x_j}(a) = \left(\dfrac{\partial f_1}{\partial x_j}(a), \ldots, \dfrac{\partial f_n}{\partial x_j}(a)\right)$.

Supondo f diferenciável no ponto a, para todo $v \in \mathbb{R}^m$ e qualquer $t \in \mathbb{R}$ suficientemente pequeno tem-se

$$f(a+tv) - f(a) = T \cdot tv + \rho(tv) \cdot |tv|, \text{ com } \lim_{v \to 0} \rho(tv) = 0.$$

Como $T \cdot tv = t \cdot T \cdot v$ e $|tv| = |t| \cdot |v|$, segue-se, para $t \neq 0$:

$$\frac{f(a+tv) - f(a)}{t} = T \cdot v \pm \rho(tv) \cdot |v|,$$

donde $\lim_{t \to 0} \dfrac{f(a+tv) - f(a)}{t} = T \cdot v$.

Portanto $T \cdot v = \dfrac{\partial f}{\partial v}(a)$.

Em particular, vemos que é única a transformação linear T que fornece a boa aproximação para o acréscimo $f(a+v) - f(a)$ na vizinhança do ponto a. Ela é chamada a *derivada* de f no ponto a e indicada com a notação $f'(a)$.

Portanto, se $f\colon U \to \mathbb{R}^n$, definida no aberto $U \subset \mathbb{R}^m$, é diferenciável no ponto $a \in U$, sua derivada é a aplicação linear $f'(a)\colon \mathbb{R}^m \to \mathbb{R}^n$, caracterizada por

$$f(a+v) - f(a) = f'(a) \cdot v + r(v), \text{ com } \lim_{v \to 0} \frac{r(v)}{|v|} = 0,$$

ou

$$f(a+v) - f(a) = f'(a) \cdot v + \rho(v)|v|, \text{ com } \lim_{v \to 0} \rho(v) = 0.$$

Às vezes usa-se também a notação $Df(a)$ em vez de $f'(a)$.

Para $n = 1$, a derivada $f'(a)$ coincide com a diferencial $df(a)$, estudada no Capítulo III. Quando $m = n = 1$, a transformação linear $f'(a)\colon \mathbb{R} \to \mathbb{R}$ confunde-se com o número $f'(a) \cdot 1$ e, para todo $v \in \mathbb{R}$, $f'(a) \cdot v$ é simplesmente o produto do número $f'(a)$ (já conhecido desde o Volume 1) pelo número v.

A transformação linear $f'(a)\colon \mathbb{R}^m \to \mathbb{R}^n$ possui, em relação às bases canônicas de $\mathbb{R}^m$ e $\mathbb{R}^n$, uma matriz $n \times m$ chamada a *matriz jacobiana* de f no ponto a, indicada com a notação $Jf(a)$. Suas m colunas são os vetores $f'(a) \cdot e_j = \dfrac{\partial f}{\partial x_j}(a) = \left(\dfrac{\partial f_1}{\partial x_j}(a), \ldots, \dfrac{\partial f_n}{\partial x_j}(a)\right)$. Assim, $Jf(a) = \left(\dfrac{\partial f_i}{\partial x_j}(a)\right)$, onde $f_1, \ldots, f_n\colon U \to \mathbb{R}$ são as funções-coordenada

de f. Cada uma das n linhas $\left(\frac{\partial f_i}{\partial x_1}(a), \ldots, \frac{\partial f_i}{\partial x_m}(a)\right)$ é, como sabemos, a matriz $1 \times m$ do funcional linear $df_i(a) =$ diferencial da i-ésima função-coordenada f_i. Para todo $v \in \mathbb{R}^m$, temos

$$f'(a) \cdot v = \frac{\partial f}{\partial v}(a) = \left(\frac{\partial f_1}{\partial v}(a), \ldots, \frac{\partial f_n}{\partial v}(a)\right) =$$
$$= (df_1(a) \cdot v, \ldots, df_n(a) \cdot v).$$

A igualdade vetorial $f(a+v) - f(a) = f'(a) \cdot v + r(v)$ equivale às n igualdades numéricas $f_i(a+v) - f_i(a) = df_i(a) \cdot v + r_i(v)$, onde $r(v) = (r_1(v), \ldots, r_n(v))$, enquanto que o limite vetorial $\lim\limits_{v \to 0} r(v)/|v| = 0$ corresponde aos n limites numéricos $\lim\limits_{v \to 0} r_i(v)/|v| = 0$. Isto prova o

Teorema 1. *A aplicação $f\colon U \to \mathbb{R}^n$ é diferenciável no ponto $a \in U$ se, e somente se, cada uma das suas funções-coordenada $f_1, \ldots, f_n\colon U \to \mathbb{R}$ é diferenciável nesse ponto.*

Corolário. *A aplicação $f = (g, h)\colon U \to \mathbb{R}^n \times \mathbb{R}^p$, dada por $f(x) = (g(x), h(x))$, é diferenciável no ponto $a \in U$ se, e somente se, cada uma das aplicações coordenadas $g\colon U \to \mathbb{R}^n$ e $h\colon U \to \mathbb{R}^p$ o é. No caso afirmativo, $f'(a) = (g'(a), h'(a))\colon \mathbb{R}^m \to \mathbb{R}^n \times \mathbb{R}^p$.*

Com efeito, as funções-coordenada de f são as funções-coordenada de g mais as de h.

O Teorema 1 permitirá, mais adiante, reduzir algumas proposições sobre aplicações diferenciáveis a fatos análogos sobre funções, já demonstrados no Capítulo III. No momento, faremos uma observação simples, dele decorrente. Se $f\colon U \to \mathbb{R}^n$, definida no aberto $U \subset \mathbb{R}^m$, é diferenciável no ponto $a \in U$ então, dado qualquer vetor $v \in \mathbb{R}^m$, temos $f'(a) \cdot v = \frac{\partial f}{\partial v}(a) = (f \circ \lambda)'(0) =$ vetor-velocidade, no instante $t = 0$, do caminho $f \circ \lambda\colon (-\delta, \delta) \to \mathbb{R}^n$, onde $\lambda\colon (-\delta, \delta) \to U$ não precisa ser o caminho retilíneo $\lambda(t) = a + tv$, mas pode ser qualquer caminho diferenciável no ponto $t = 0$, tal que $\lambda(0) = a$ e $\lambda'(0) = v$. Basta notar que $\frac{\partial f}{\partial v}(a) = \left(\frac{\partial f_1}{\partial v}(a), \ldots, \frac{\partial f_n}{\partial v}(a)\right)$ e este fato é verdadeiro para funções diferenciáveis. (Veja a observação após o Corolário 2 da Regra da Cadeia, no §3 do Capítulo III.)

Uma aplicação $f\colon U \to \mathbb{R}^n$ diz-se *diferenciável* no aberto $U \subset \mathbb{R}^m$ quando é diferenciável em todos os pontos de U. Neste caso, fica definida

a *aplicação derivada* $f'\colon U \to \mathcal{L}(\mathbb{R}^m;\mathbb{R}^n)$, que associa a cada ponto $x \in U$ a transformação linear $f'(x)\colon \mathbb{R}^m \to \mathbb{R}^n$, derivada de f naquele ponto. Fica também definida, para cada $v \in \mathbb{R}^m$, a aplicação $\dfrac{\partial f}{\partial v}\colon U \to \mathbb{R}^n$, cujo valor num ponto $x \in U$ é a derivada direcional $\dfrac{\partial f}{\partial v}(x) = f'(x)\cdot v$.

O espaço vetorial $\mathcal{L}(\mathbb{R}^m;\mathbb{R}^n)$, formado pelas transformações lineares $T\colon \mathbb{R}^m \to \mathbb{R}^n$, possui, como sabemos, uma norma natural, dada por $|T| = \sup\{|T\cdot v|; v \in \mathbb{R}^m, |v| = 1\}$. Identificaremos $\mathcal{L}(\mathbb{R}^m;\mathbb{R}^n)$ com $\mathbb{R}^{nm}$ fazendo corresponder a cada transformação linear $T\colon \mathbb{R}^m \to \mathbb{R}^n$ sua matriz em relação às bases canônicas de $\mathbb{R}^m$ e $\mathbb{R}^n$. Com esta identificação, as funções-coordenada de uma aplicação $\varphi\colon X \to \mathcal{L}(\mathbb{R}^m;\mathbb{R}^n)$, definida num conjunto $X \subset \mathbb{R}^p$, são as nm funções $\varphi_{ij}\colon X \to \mathbb{R}$ tais que, para todo $x \in X$, $\varphi_{ij}(x) = (i,j)$-ésimo elemento da matriz da transformação linear $\varphi(x)$.

Como duas normas quaisquer em $\mathbb{R}^{nm}$ são equivalentes, resulta então do Teorema 12, Capítulo I, que uma aplicação $\varphi\colon X \to \mathcal{L}(\mathbb{R}^m;\mathbb{R}^n)$ é contínua se, e somente se, cada elemento da matriz de $\varphi(x)$ é uma função contínua de x. Resulta, por sua vez, do Teorema 1 acima que uma aplicação $\varphi\colon U \to \mathcal{L}(\mathbb{R}^m;\mathbb{R}^n)$ é diferenciável no ponto $a \in U$ se, e somente se, cada uma das funções $\varphi_{ij}\colon I \to \mathbb{R}$, dadas por $\varphi_{ij}(x) = (i,j)$-ésimo elemento da matriz de $\varphi(x)$, é diferenciável no ponto a. Estas considerações levam ao

Teorema 2. *Seja $f\colon U \to \mathbb{R}^n$ definida no aberto $U \subset \mathbb{R}^m$. As seguintes afirmações são equivalentes:*

(1) *f é diferenciável e a aplicação derivada $f'\colon U \to \mathcal{L}(\mathbb{R}^m;\mathbb{R}^n)$ é contínua;*

(2) *As funções-coordenada $f_1,\dots,f_n\colon U \to \mathbb{R}$ da aplicação f possuem derivadas parciais contínuas $\dfrac{\partial f_i}{\partial x_j}\colon U \to \mathbb{R}$;*

(3) *Para cada $v \in \mathbb{R}^m$, existe a derivada direcional $(\partial f/\partial v)(x)$ em qualquer ponto $x \in U$ e a aplicação $\partial f/\partial v\colon U \to \mathbb{R}^n$ é contínua.*

Demonstração: (1) $\Rightarrow$ (2) porque as derivadas parciais $\partial f_i/\partial x_j$ são as funções-coordenada da alicação f'. Pelo Teorema 1, Capítulo III, (2) implica que cada função f_i é diferenciável, logo f é diferenciável (Teorema 1, acima) e a aplicação derivada f' é contínua pois suas funções-coordenada o são. Isto mostra que (2) $\Rightarrow$ (1). Em seguida, supondo (2)

temos f diferenciável, logo $\dfrac{\partial f}{\partial v} = \Sigma\, \alpha_j \,\dfrac{\partial f}{\partial x_j}$, onde $v = (\alpha_1, \ldots, \alpha_m)$. Ora, cada aplicação $\dfrac{\partial f}{\partial x_j} \colon U \to \mathbb{R}^n$ é contínua porque suas funções-coordenada o são. Logo $\dfrac{\partial f}{\partial v}$ é contínua, isto é, (2) $\Rightarrow$ (3). Finalmente, supondo (3) então, em particular, tomando $v = e_j$ vemos que para $j = 1, \ldots, m$, as derivadas parciais $\partial f/\partial x_j \colon U \to \mathbb{R}^n$ são contínuas, logo é contínua cada uma das suas funções-coordenada $\partial f_i/\partial x_j \colon U \to \mathbb{R}$. Assim, (3) $\Rightarrow$ (2), o que completa a demonstração.

Diz-se que a aplicação $f \colon U \to \mathbb{R}^n$ é de *classe* C^1 no aberto $U \subset \mathbb{R}^m$, e escreve-se $f \in C^1$, para significar que f cumpre uma das (e portanto todas as) condições do teorema acima.

Em particular, $f \in C^1$ se, e somente se, cada uma das suas funções-coordenada é de classe C^1.

Uma aplicação $f \colon U \to \mathbb{R}^n$, definida no aberto $U \subset \mathbb{R}^m$, diz-se *duas vezes diferenciável* no ponto $a \in U$ quando é diferenciável em U e cumpre as condições abaixo:

(1) A aplicação derivada $f' \colon U \to \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$ é diferenciável no ponto a;

(2) Cada derivada parcial $\dfrac{\partial f_i}{\partial x_j} \colon U \to \mathbb{R}$ é uma função diferenciável no ponto a;

(3) Para cada vetor $v \in \mathbb{R}^m$, a derivada direcional $\partial f/\partial v \colon U \to \mathbb{R}^n$ é uma aplicação diferenciável no ponto a.

Como no Teorema 2, vê-se facilmente que as três condições acima são equivalentes, de modo que f cumpre uma delas se, e somente se, cumpre todas.

Quando $f \colon U \to \mathbb{R}^n$ é duas vezes diferenciável no ponto $a \in U \subset \mathbb{R}^m$, sua *derivada segunda* no ponto a é a aplicação bilinear

$$f''(a) \colon \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^n,$$

cujo valor no ponto $(v, w) \in \mathbb{R}^m \times \mathbb{R}^m$ é o vetor

$$f''(a) \cdot v \cdot w = \frac{\partial}{\partial w} \left(\frac{\partial f}{\partial v} \right) (a) \in \mathbb{R}^n.$$

Como é natural, escreveremos $\dfrac{\partial^2 f}{\partial w \partial v}(a)$ em vez de $\dfrac{\partial}{\partial w}\left(\dfrac{\partial f}{\partial v}\right)(a)$. Se $v = (\alpha_1, \ldots, \alpha_m)$ e $w = (\beta_1, \ldots, \beta_m)$ então

$$f''(a) \cdot v \cdot w = \frac{\partial}{\partial w}\left(\frac{\partial f}{\partial v}\right)(a) = \frac{\partial}{\partial w}\left(\sum_j \alpha_j \frac{\partial f}{\partial x_j}\right)(a) =$$

$$= \sum_{j,k=1}^{m} \alpha_j \beta_k \frac{\partial^2 f}{\partial x_k \partial x_j}(a)$$

é o vetor de $\mathbb{R}^n$ cujas coordenadas são os números

$$f_i''(a) \cdot v \cdot w = \sum_{j,k=1}^{m} \frac{\partial^2 f_i}{\partial x_k \partial x_j}(a) \cdot \alpha_j \beta_k\,,\ i = 1, \ldots, n.$$

Segue-se do Teorema de Schwarz para funções (Capítulo III, §7) que $f_i''(a) \cdot v \cdot w = f_i''(a) \cdot w \cdot v$ para todo $i \in [1, n]$, donde $f''(a) \cdot v \cdot w = f''(a) \cdot w \cdot v$, ou seja $\dfrac{\partial^2 f}{\partial w \partial v}(a) = \dfrac{\partial^2 f}{\partial v \partial w}(a)$ quando a aplicação $f \colon U \to \mathbb{R}^n$ é duas vezes diferenciável no ponto a. Isto nos dá o

Teorema de Schwarz (para aplicações). *Se a aplicação $f \colon U \to \mathbb{R}^n$, definida no aberto $U \subset \mathbb{R}^m$, é duas vezes diferenciável no ponto $a \in U$, então a derivada segunda $f''(a) \colon \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^n$ é uma aplicação bilinear simétrica.*

Uma aplicação $f \colon U \to \mathbb{R}^n$ diz-se de *classe* C^2 no aberto $U \subset \mathbb{R}^m$ quando é diferenciável e sua derivada $f' \colon U \to \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$ é de classe C^1. Isto equivale a dizer que, para cada $v \in \mathbb{R}^m$, a derivada $\partial f / \partial v \colon U \to \mathbb{R}^n$ é de classe C^1 ou ainda, que, para $i \in [1, n]$ e $j, k \in [1, m]$ arbitrários, existem e são contínuas as derivadas segundas $\partial^2 f_i / \partial x_j \partial x_k \colon U \to \mathbb{R}$ das funções-coordenada de f.

Por indução, dizemos que a aplicação $f \colon U \to \mathbb{R}^n$, definida no aberto $U \subset \mathbb{R}^m$, é k vezes diferenciável no ponto $a \in U$ quando f é diferenciável em U, e a aplicação derivada $f' \colon U \to \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$ é $k - 1$ vezes diferenciável no ponto a, o que equivale a dizer que, para todo vetor $v \in \mathbb{R}^m$, a derivada direcional $\partial f / \partial v \colon U \to \mathbb{R}^n$ é uma aplicação $k-1$ vezes diferenciável no ponto a, ou ainda, que as derivadas parciais $\partial f_i / \partial x_j \colon U \to \mathbb{R}$ são todas funções $k - 1$ vezes diferenciáveis no ponto a.

Quando $f \colon U \to \mathbb{R}^n$ é k vezes diferenciável no ponto a, define-se a *k-ésima derivada* (ou *derivada de ordem k*) de f no ponto a como a

aplicação k-linear

$$f^{(k)}(a)\colon \mathbb{R}^m \times \cdots \times \mathbb{R}^m \to \mathbb{R}^n$$

cujo valor no ponto $(v_1, \ldots, v_k) \in \mathbb{R}^m \times \cdots \times \mathbb{R}^m$ é o vetor

$$f^{(k)}(a) \cdot v_1 \cdots v_k = \frac{\partial^k f}{\partial v_k \partial v_{k-1} \ldots \partial v_1}(a) \in \mathbb{R}^n.$$

É um corolário do Teorema de Schwarz que a k-ésima derivada $f^{(k)}(a)$ é uma aplicação k-linear simétrica. Por exemplo, se $k = 3$ e $n = 1$, dados os vetores $u = (\alpha_1, \ldots, \alpha_m)$, $v = (\beta_1, \ldots, \beta_m)$ e $w = (\gamma_1, \ldots, \gamma_m)$, temos

$$f^{(3)}(a) \cdot u \cdot v \cdot w = \sum \frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k}(a)\alpha_i \beta_j \gamma_k$$

e sabemos do Capítulo III que as derivadas parciais $\partial^3 f / \partial x_i \partial x_j \partial x_k$ independem da ordem dos índices i, j, k.

A aplicação $f\colon U \to \mathbb{R}^n$ diz-se *de classe* C^k no aberto $U \subset \mathbb{R}^m$ quando é diferenciável e sua derivada $f'\colon U \to \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$ é uma aplicação de classe C^{k-1}. Isto equivale a dizer que, para todo $v \in \mathbb{R}^m$, a aplicação $\partial f / \partial v\colon U \to \mathbb{R}^n$ é de classe C^{k-1}, ou que existem e são contínuas em U todas as derivadas parciais de ordem $\leq k$ das funções coordenadas de f. Escreve-se, então $f \in C^k$.

Para completar, f diz-se *de classe* C^0 quando é contínua, e de classe C^∞ quando $f \in C^k$ para todo $k = 0, 1, 2, \ldots$ Evidentemente, $f \in C^k \Rightarrow f \in C^{k-1}$.

Quando $v_1 = \cdots = v_k = v$, o valor da aplicação k-linear $f^{(k)}(a)$ na k lista $(v, \ldots, v)$ será indicado com $f^{(k)}(a) \cdot v \ldots v$. Em particular, se $f\colon U \to \mathbb{R}$ é uma *função* k vezes diferenciável no ponto $a \in U$, sua k-ésima diferencial no ponto a, definida no Capítulo III é $d^k(a) \cdot v^k = f^{(k)}(a) \cdot v^k$.

Finalmente, observamos que $f = (g, h)\colon U \to \mathbb{R}^n \times \mathbb{R}^p$, dada por $f(x) = (g(x), h(x))$, é k vezes diferenciável num ponto (ou de classe C^k em U) se, e somente se, suas aplicações coordenadas $g\colon U \to \mathbb{R}^n$, $h\colon U \to \mathbb{R}^p$ são k vezes diferenciáveis nesse ponto (ou de classe C^k em U).

2 Exemplos de aplicações diferenciáveis

Exemplo 1. Toda aplicação constante é de classe C^∞ e sua derivada é nula.

Exemplo 2. Toda aplicação linear $T\colon \mathbb{R}^m \to \mathbb{R}^n$ é diferenciável e, em cada ponto $x \in \mathbb{R}^m$, temos $T'(x) = T$. Com efeito, $T(x+v) - T(x) = T \cdot v + 0$. A aplicação derivada $T'\colon \mathbb{R}^m \to \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$ é constante, logo $T \in C^\infty$.

Exemplo 3. Toda aplicação bilinear $\varphi\colon \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^p$ é diferenciável e, em cada ponto $(a,b) \in \mathbb{R}^m \times \mathbb{R}^n$, sua derivada é a transformação linear $\varphi'(a,b)\colon \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^p$, definida por

$$\varphi'(a,b) \cdot (v,w) = \varphi(v,b) + \varphi(a,w).$$

Com efeito, a bilinearidade de φ nos dá

$$\varphi(a+v, b+w) - \varphi(a,b) = \varphi(v,b) + \varphi(a,w) + \varphi(v,w).$$

Sabemos (veja a observação (7.2) no §7 do Capítulo I) que existe uma constante $c > 0$ tal que $|\varphi(v,w)| \le c|v|\,|w|$ para todo $v \in \mathbb{R}^m$ e todo $w \in \mathbb{R}^n$. Fazendo uso da norma da soma, temos $|(v,w)| = |v| + |w|$. Logo

$$\frac{|\varphi(v,w)|}{|(v,w)|} \le \frac{c|v|\,|w|}{|v|+|w|} \le c|v| \text{ e daí, } \lim_{(v,w)\to(0,0)} \frac{\varphi(v,w)}{|(v,w)|} = 0.$$

Fica, então, cumprida a condição de diferenciabilidade, com $\varphi'(a,b)\colon (v,w) \mapsto \varphi(v,b) + \varphi(a,w)$ e $r(v,w) = \varphi(v,w)$.

O mesmo raciocínio mostra que se $\xi\colon \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}^k$ é trilinear então ξ é diferenciável em todo ponto $(a,b,c) \in \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^p$, sendo sua derivada $\xi'(a,b,c)\colon \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^k$ a aplicação linear dada por

$$\xi'(a,b,c) \cdot (u,v,w) = \xi(u,b,c) + \xi(a,v,c) + \xi(a,b,w).$$

Mais geralmente, toda aplicação k-linear $\mu\colon \mathbb{R}^{m_1} \times \cdots \times \mathbb{R}^{m_k} \to \mathbb{R}^n$ é diferenciável e sua derivada no ponto $a = (a_1, \ldots, a_k)$ é a transformação linear $\mu'(a)\colon \mathbb{R}^{m_1} \times \cdots \times \mathbb{R}^{m_k} \to \mathbb{R}^n$, dada por

$$\mu'(a) \cdot (v_1, \ldots, v_k) = \sum_{i=1}^{k} \mu(a_1, \ldots, a_{i-1}, v_i, a_{i+1}, \ldots, a_k).$$

Casos particulares importantes de aplicações bilineares são o produto interno $\varphi\colon \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$, $\varphi(x,y) = \langle x,y\rangle$, e a multiplicação de matrizes $\mu\colon \mathbb{R}^{pn} \times \mathbb{R}^{nm} \to \mathbb{R}^{pm}$, $\mu(X,Y) = XY$, cujas derivadas são dadas por $\varphi'(x,y)\cdot(v,w) = \langle v,y\rangle + \langle x,w\rangle$ e $\mu'(X,Y)\cdot(V,W) = VY + XW$.

Um exemplo interessante de aplicação n-linear é a função determinante, $\det : \mathbb{R}^n \times \cdots \times \mathbb{R}^n = \mathbb{R}^{n^2} \to \mathbb{R}$, onde escreveremos $\det X = \det(X_1,\ldots,X_n)$ para enfatizar que $\det X$ depende linearmente de cada uma das n linhas X_i da matriz $X = (X_1,\ldots,X_n)$. Sua derivada no ponto X é o funcional linear $\det{}'(X)\colon \mathbb{R}^{n^2} \to \mathbb{R}$, cujo valor a matriz $V = (V_1,\ldots,V_n)$ é

$$\det{}'(X)\cdot V = \sum_{k=1}^{n} \det(X_1,\ldots,X_{k-1},V_k,X_{k+1},\ldots,X_n).$$

Em particular, se $V = E_{ij}$ = matriz cujo (i,j)-ésimo elemento é igual a 1 e os demais são zero, então todas as parcelas do somatório acima são nulas, salvo a i-ésima, que é igual a $(-1)^{i+j}\,X_{[i,j]}$, onde $X_{[i,j]}$ é o determinante da matriz $(n-1)\times(n-1)$ obtida de X pela omissão da i-ésima linha e da j-ésima coluna. Então

$$\frac{\partial \det}{\partial x_{ij}}(X) = \det{}'(X)\cdot E_{ij} = (-1)^{i+j}\,X_{[i,j]}\,,$$

e reobtemos assim um fato já conhecido. (Exemplo 26 do Capítulo III.)

Toda aplicação k-linear é de classe C^∞ pois suas funções-coordenada são ainda k-lineares e portanto são polinômios. Por exemplo, se $f\colon \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$ é uma função bilinear, pomos $f(e_i,e_j) = a_{ij}$ e, para quaisquer $x = (x_1,\ldots,x_m)$, $y = (y_1,\ldots,y_n)$ temos $f(x,y) = \sum_{i,j} a_{ij}x_iy_j$. Ou ainda: se $f\colon \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}$ é trilinear, pondo $f(e_i,e_j,e_k) = a_{ijk}$ vem $f(x,y,z) = \sum_{i,j,k} a_{ijk}\,x_iy_jz_k$.

Exemplo 4. O conjunto $U = \mathrm{GL}(\mathbb{R}^n) \subset \mathbb{R}^{n^2}$, formado pelas matrizes invertíveis $n\times n$, é aberto porque a função determinante é contínua e $X \in U \Leftrightarrow \det X \neq 0$. Mostraremos que a aplicação $f\colon U \to \mathbb{R}^{n^2}$, dada por $f(X) = X^{-1}$, é diferenciável e sua derivada no ponto $A \in U$ é a transformação linear $f'(A)\colon \mathbb{R}^{n^2} \to \mathbb{R}^{n^2}$, definida por $f'(A)\cdot V = -A^{-1}\cdot V\cdot A^{-1}$. Com efeito, se escrevermos

$$(A+V)^{-1} - A^{-1} = -A^{-1}\cdot V\cdot A^{-1} + r(V)$$

e multiplicarmos ambos os membros desta igualdade, à esquerda, por $A+V$, obteremos

$$I - I - VA^{-1} = -VA^{-1} - (VA^{-1})^2 + (A+V)\cdot r(V),$$

donde

$$r(V) = (A+V)^{-1}\cdot (VA^{-1})^2.$$

[Note que, sendo U aberto, existe $(A+V)^{-1}$ para todo V com norma suficientemente pequena.] Daí resulta

$$|r(V)| \le |(A+V)^{-1}|\cdot |A^{-1}|^2\cdot |V|^2.$$

Segue-se então do lema abaixo que $\lim\limits_{V\to 0} r(V)/|V| = 0$.

Lema. *Seja $A\in\mathbb{R}^{n^2}$ invertível. Existe $c>0$ tal que, para toda matriz $n\times n$, V, com $|V|\le c$, $A+V$ é invertível e $|(A+V)^{-1}|\le 1/c$.*

Demonstração: Tomando $c = (2|A^{-1}|)^{-1}$ vem $|x| = |A^{-1}Ax| \le |A^{-1}|\,|Ax|$, donde $|Ax|\ge 2c|x|$ para todo $x\in\mathbb{R}^n$. Então $|V|\le c \Rightarrow |(A+V)x| = |Ax+Vx| \ge |Ax| - |Vx| \ge 2c|x| - c|x| = c|x|$. Logo $|V|\le c\Rightarrow A+V$ invertível e $|x| = |(A+V)(A+V)^{-1}x| \ge c|(A+V)^{-1}x|$, donde $|(A+V)^{-1}x| \le |x|/c$ para todo $x\in\mathbb{R}^n$, ou seja, $|V|\le c\Rightarrow |(A+V)^{-1}|\le 1/c$.

Em particular, a inversão de matrizes $f\colon X\to X^{-1}$ é uma aplicação contínua de U em $\mathbb{R}^{n^2}$. Como $f(U)=U$ e $f=f^{-1}$, vemos que f é um homeomorfismo de U sobre si mesmo.

Completando o Exemplo 4, mostraremos agora que a inversão de matrizes $f\colon U\to\mathbb{R}^{n^2}$ é uma aplicação C^∞. Para isso, fixemos $V\in\mathbb{R}^{n^2}$. A derivada direcional $\partial f/\partial V\colon U\to\mathbb{R}^{n^2}$ é dada por $\partial f/\partial V(X) = -X^{-1}VX^{-1}$. Seja $\varphi\colon \mathbb{R}^{n^2}\times\mathbb{R}^{n^2}\to\mathbb{R}^{n^2}$ a aplicação bilinear tal que $\varphi(X,Y) = XVY$. Então $\partial f/\partial V = -\varphi\circ(f,f)$, onde $(f,f)\colon U\to\mathbb{R}^{n^2}\times\mathbb{R}^{n^2}$ é dada por $(f,f)(X) = (X^{-1},X^{-1})$.

$$U \xrightarrow{(f,f)} \mathbb{R}^{n^2}\times\mathbb{R}^{n^2} \xrightarrow{-\varphi} \mathbb{R}^{n^2}$$

$$X\mapsto (X^{-1},X^{-1})\mapsto -X^{-1}VX^{-1} = \frac{\partial f}{\partial V}(S).$$

Segue-se que $\partial f/\partial V$ é contínua, seja qual for V, donde $f\in C^1$. Então, $(f,f)\in C^1$. Pelo Exemplo 3, a aplicação bilinear φ é de classe C^∞. Admitindo (como será mostrado no §3, logo a seguir) que a composta de

duas aplicações de classe C^k é de classe C^k, concluiremos que $\partial f/\partial V = \varphi \circ (f, f)$ é de classe C^1, para todo $V \in \mathbb{R}^{n^2}$. Logo $f \in C^2$. Repetindo o argumento, veremos que $f \in C^k$ para todo k, ou seja $f \in C^\infty$.

Sejam U, V abertos do espaço euclidiano. Uma bijeção $f\colon U \to V$ chama-se um *difeomorfismo* de U sobre V quando é diferenciável e sua inversa $f^{-1}\colon V \to U$ também o é.

A inversão de matrizes é um exemplo de difeomorfismo $f\colon U \to U$, de classe C^∞, onde U é o conjunto das matrizes invertíveis $n \times n$. Neste caso particular, temos $f^{-1} = f$.

Observação: O leitor, com alguma razão, pode indagar como se atinou com a forma da derivada $f'(A)\colon V \mapsto -A^{-1} \cdot V \cdot A^{-1}$ para a inversão de matrizes $f\colon X \mapsto X^{-1}$. Esta é uma boa pergunta, em geral. Existem critérios indiretos, como o Teorema das Funções Implícitas, (ou, em alguns casos, o Teorema 2 acima) que permitem concluir que uma certa aplicação é diferenciável, sem que se conheça como opera sua derivada. Na ausência desses métodos indiretos, fica o problema de obter um candidato razoável para a derivada, sem o qual não se pode provar a diferenciabilidade da aplicação. Um processo aceitável (quando pode ser aplicado) é o de desenvolver $f(a+v)$ (ou cada uma das suas funções-coordenada) em série de potências nas coordenadas de v, e destacar a parte do primeiro grau em relação a v: esta será $f'(a) \cdot v$. No caso em questão, temos $f(A+V) = (A+V)^{-1} = [(I + VA^{-1}) \cdot A]^{-1} = A^{-1}(I + VA^{-1})^{-1}$. Ora, a identidade $(I-X)(I+X+\cdots+X^n) = I - X^{n+1}$, onde $X \in \mathbb{R}^{m^2}$, mostra que, quando $|X| < 1$ (caso em que $I - X$ é obviamente injetiva, donde invertível), fazendo $n \to \infty$, tem-se

$$(I - X) \cdot \sum_{n=0}^{\infty} X^n = I, \text{ donde } (I - X)^{-1} = \sum_{n=0}^{\infty} X^n.$$

Tomando $X = -VA^{-1}$, concluimos que, quando $|V| < |A^{-1}|^{-1}$, vale

$$(I + VA^{-1})^{-1} = \sum_{n=0}^{\infty} (-1)^n (VA^{-1})^n,$$

donde

$$(A + V)^{-1} = A^{-1} \cdot \sum_{n=0}^{\infty} (-1)^n (VA^{-1})^n = A^{-1} - A^{-1} \cdot V \cdot A^{-1} + r(V),$$

onde

$$r(V) = \sum_{n=2}^{\infty} (-1)^n A^{-1}(VA^{-1})^n$$

cumpre a condição $\lim_{V \to 0} r(V)/|V| = 0$. Segue-se que $f'(A) \cdot V = -A^{-1} \cdot V \cdot A^{-1}$. Note-se ainda que, quando $n = 1$, esta fórmula diz simplesmente que se $f(x) = 1/x$ então $f'(a) = -1/a^2$.

Exemplo 5. Uma função de variável complexa $f\colon U \to \mathbb{C}$, definida no aberto $U \subset \mathbb{C}$, pode ser vista como uma aplicação $f\colon U \to \mathbb{R}^2$, definida no aberto $U \subset \mathbb{R}^2$. A derivada da função complexa f no ponto $z = x+iy$ é o número complexo definido pelo limite

$$f'(z) = \lim_{H \to 0} \frac{f(z+H) - f(z)}{H}\,,$$

quando tal limite existe. Isto equivale a afirmar que

$$f(z+H) - f(z) = f'(z) \cdot H + r(H), \text{ onde } \lim_{H \to 0} \frac{r(H)}{|H|} = 0.$$

Acima, $f'(z) \cdot H$ é uma multiplicação de números complexos. Portanto, a função complexa $f\colon U \to \mathbb{C}$ é derivável no ponto $z \in U$ se, e somente se, a aplicação $f\colon U \to \mathbb{R}^2$ é diferenciável nesse ponto e sua derivada $f'(z)\colon \mathbb{R}^2 \to \mathbb{R}^2$ é uma transformação linear do plano que consiste em multiplicar por um número complexo fixo. Ora, as transformações deste tipo são: a transformação nula (que consiste em multiplicar pelo número 0) e as semelhanças positivas. (Veja o §3 do Capítulo I.) Se $f = u + iv$, as equações de Cauchy-Riemann $\frac{\partial u}{\partial x}(z) = \frac{\partial v}{\partial y}(z)$ e $\frac{\partial u}{\partial y}(z) = -\frac{\partial v}{\partial x}(z)$, vistas no Exemplo 6 do Capítulo III, exprimem meramente o fato de que a matriz jacobiana $Jf(z)$ tem a forma $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$, como toda matriz de uma semelhança positiva, que consiste na multiplicação por $a+ib$. Vistas de outro ângulo, as equações de Cauchy-Riemann dizem que o vetor $f'(z) \cdot e_2 = \left(\frac{\partial u}{\partial y}(z), \frac{\partial v}{\partial y}(z)\right)$ é obtido de $f'(z) \cdot e_1 = \left(\frac{\partial u}{\partial x}(z), \frac{\partial v}{\partial x}(z)\right)$ por meio de uma rotação de 90° no sentido anti-horário (multiplicação por i). Como e_2 é obtido de e_1 da mesma forma, isto caracteriza o fato de que a transformação linear $f'(z)\colon \mathbb{R}^2 \to \mathbb{R}^2$, ou é nula ou preserva ângulos orientados.

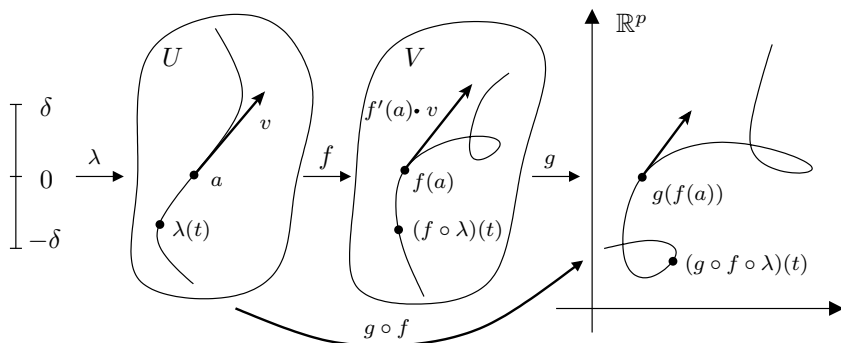
Exemplo 6. Seja $f\colon I \to \mathbb{R}^n$ um caminho definido no intervalo aberto $I \subset \mathbb{R}$. No Capítulo II, dissemos que f é diferenciável no ponto $a \in I$ quando existe o vetor-velocidade $v = \lim\limits_{t\to 0}[f(a+t)-f(a)]/t$. Isto equivale a escrever $f(a+t) - f(a) = t \cdot v + r(t)$, onde $\lim\limits_{t\to 0} r(t)/t = 0$. Ora, existe um isomorfismo natural $\mathcal{L}(\mathbb{R}, \mathbb{R}^n) \approx \mathbb{R}^n$, que faz corresponder a cada transformação linear $T\colon \mathbb{R} \to \mathbb{R}^n$ o vetor $v = T \cdot 1$ e, reciprocamente, a cada vetor $v \in \mathbb{R}^n$ associa a transformação linear $T\colon \mathbb{R} \to \mathbb{R}^n$ com $T \cdot t = t \cdot v$. Mediante este isomorfismo, um caminho é diferenciável no sentido do Capítulo 2 se, e somente se, é diferenciável no sentido deste capítulo.

3 A regra da cadeia

A versão intrínseca da Regra da Cadeia, que apresentaremos neste parágrafo, diz essencialmente que a derivada da aplicação composta é a composta das derivadas (tomadas nos pontos adequados).

Regra da Cadeia (para aplicações). *Sejam $U \subset \mathbb{R}^m$, $V \subset \mathbb{R}^n$ abertos, $f\colon U \to \mathbb{R}^n$ diferenciável no ponto a, com $f(U) \subset V$, e $g\colon V \to \mathbb{R}^p$ diferenciável no ponto $f(a)$. Então $g \circ f\colon U \to \mathbb{R}^p$ é diferenciável no ponto a, com $(g \circ f)'(a) = g'(f(a)) \cdot f'(a)\colon \mathbb{R}^m \to \mathbb{R}^p$.*

Demonstração: Sejam $g_1, \ldots, g_p\colon V \to \mathbb{R}$ as funções-coordenada de g. As funções-coordenada de $g \circ f$ são $g_1 \circ f, \ldots, g_p \circ f\colon U \to \mathbb{R}$. Pelo Teorema 1, juntamente com a Regra da Cadeia para funções, elas são diferenciáveis no ponto a, logo $g \circ f$ é diferenciável nesse ponto. Seja agora $v \in \mathbb{R}^m$ um vetor arbitrário. Tomemos um caminho $\lambda\colon (-\delta, \delta) \to U$, diferenciável no ponto $t = 0$, com $\lambda(0) = a$ e $\lambda'(0) = v$. (Por exemplo, $\lambda(t) = a + tv$.) Então $f \circ \lambda$ é um caminho em V, diferenciável no ponto $t = 0$, com $(f \circ \lambda)(0) = f(a)$ e $(f \circ \lambda)'(0) = f'(a) \cdot v$.

Analogamente, $g \circ f \circ \lambda$ é um caminho em $\mathbb{R}^p$, diferenciável no instante $t = 0$ e seu vetor-velocidade é $(g \circ (f \circ \lambda))'(0) = g'(f(a)) \cdot f'(a) \cdot v$ por um lado, e, por outro lado, é $((g \circ f) \circ \lambda)'(0) = (g \circ f)'(a) \cdot v$. Segue-se que $(g \circ f)'(a) \cdot v = g'(f(a)) \cdot f'(a) \cdot v$ para todo $v \in \mathbb{R}^m$, logo $(g \circ f)'(a) = g'(f(a)) \cdot f'(a)\colon \mathbb{R}^m \to \mathbb{R}^p$.

Corolário 1. *Sejam* $Jf(a) = \left(\dfrac{\partial f_i}{\partial x_j}(a)\right)$, $Jg(f(a)) = \left(\dfrac{\partial g_i}{\partial x_j}(f(a))\right)$ *e* $J(g\circ f)(a) = \left(\dfrac{\partial(g_i\circ f)}{\partial x_j}(a)\right)$ *as matrizes jacobianas das aplicações* f, g *e* $g\circ f$ *nos pontos indicados. Supondo* f *diferenciável no ponto* a *e* g *diferenciável no ponto* $f(a)$, *tem-se* $J(g\circ f)(a) = Jg(f(a))\cdot Jf(a)$.

Observe que Jf é uma matriz $n\times m$, Jg é $p\times n$ e $J(g\circ f)$ é $p\times m$. Como as matrizes jacobianas são simplesmente as matrizes das derivadas, a igualdade $J(g\circ f)(a) = Jg(f(a))\cdot Jf(a)$ é apenas a formulação da Regra da Cadeia em termos de matrizes, em vez de transformações lineares. Pela definição bem conhecida da multiplicação de matrizes, temos, para cada $i = 1,\dots,p$ e cada $j = 1,\dots,m$:

$$\frac{\partial(g_i\circ f)}{\partial x_j}(a) = \sum_{k=1}^{n}\frac{\partial g_i}{\partial y_k}(f(a))\cdot\frac{\partial f_k}{\partial x_j}(a).$$

As igualdades acima são a Regra da Cadeia para as funções-coordenada $g_i\circ f$. Convém notar, entretanto, que a demonstração acima não permite que se conclua dela uma nova prova da Regra da Cadeia para funções, pois esta foi usada tacitamente quando admitimos que se pode obter $g'(f(a))\cdot w$ (com $w = f'(a)\cdot v$) como o vetor-velocidade do caminho $g\circ f\circ\lambda$ no ponto $t = 0$, usando apenas que $(f\circ\lambda)'(0) = w$ e $(f\circ\lambda)(0) = f(a)$.

Se f e g são diferenciáveis em todos os pontos dos seus domínios, podemos escrever a igualdade de funções abaixo, para quaisquer $i\in[1,p]$ e $j\in[1,m]$:

$$\frac{\partial(g_i\circ f)}{\partial x_j} = \sum_{k=1}^{n}\left(\frac{\partial g_i}{\partial y_k}\circ f\right)\cdot\frac{\partial f_k}{\partial x_j}\colon U\to\mathbb{R}.$$

Daí decorre (como no Corolário 2 do Teorema 1, no final do §3, Capítulo III), com auxílio do Teorema 2 acima, o

Corolário 2. *A composta de duas aplicações de classe C^k é uma aplicação de classe C^k.*

É instrutivo formular a demonstração do corolário acima em termos intrínsecos. A idéia da prova é exatamente a mesma, mudando apenas a linguagem. Sejam pois $f, g \in C^k$. A Regra da Cadeia diz que, para cada $x \in U$, $(g \circ f)'(x) = g'(f(x)) \cdot f'(x)$. Considerando as aplicações derivadas $f' \colon U \to \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$, $g' \colon V \to \mathcal{L}(\mathbb{R}^n; \mathbb{R}^p)$ e $(g \circ f)' \colon U \to \mathcal{L}(\mathbb{R}^m; \mathbb{R}^p)$, a igualdade acima se traduz em:

$$(g \circ f)' = (g' \circ f) \cdot f' \colon U \to \mathcal{L}(\mathbb{R}^m; \mathbb{R}^p),$$

onde o símbolo $\circ$ indica composição de aplicações e $\cdot$ significa produto de transformações lineares. Podemos ainda considerar a multiplicação de transformações lineares como uma aplicação bilinear $\mu \colon \mathcal{L}(\mathbb{R}^n; \mathbb{R}^p) \times \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n) \to \mathcal{L}(\mathbb{R}^m; \mathbb{R}^p)$, onde $\mu(T, S) = T \cdot S$. Então a Regra da Cadeia para as aplicações derivadas se exprime, em termos ainda mais explícitos, como

$$(*) \qquad (g \circ f)' = \mu \circ (g' \circ f, f'),$$

onde $(g' \circ f, f') \colon U \to \mathcal{L}(\mathbb{R}^n; \mathbb{R}^p) \times \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$ é a aplicação que tem por coordenadas $g' \circ f$ e f'. Examinemos a igualdade (*). Sabemos que $\mu \in C^\infty$, isto é, $\mu \in C^k$ para todo k. Começamos supondo $f, g \in C^1$. Então $g', f' \in C^0$, por definição. Logo $g' \circ f, f' \in C^0$, donde $\mu \circ (g' \circ f, f') \in C^0$ e daí $(g \circ f)' \in C^0$ o que significa $g \circ f \in C^1$. Se, porém $g, f \in C^2$ então, por definição, $g', f' \in C^1$. Pelo que acabamos de ver, isto implica $g' \circ f \in C^1$. Logo $(g' \circ f, f') \in C^1$ e, como $\mu \in C^1$, $\mu \circ (g' \circ f, f') \in C^1$, isto é, $(g \circ f)' \in C^1$, donde $g \circ f \in C^2$. E assim por diante.

Corolário 3. *Se a aplicação $f \colon U \to \mathbb{R}^n$, definida no aberto $U \subset \mathbb{R}^m$ e diferenciável no ponto $a \in U$, admite uma inversa $g = f^{-1} \colon V \to \mathbb{R}^m$, definida no aberto $V \subset \mathbb{R}^n$ e diferenciável no ponto $b = f(a)$, então $f'(a) \colon \mathbb{R}^m \to \mathbb{R}^n$ é um isomorfismo, cujo inverso é $g'(b) \colon \mathbb{R}^n \to \mathbb{R}^m$. Em particular, $m = n$.*

De $g \circ f = \mathrm{id}_U$ e $f \circ g = \mathrm{id}_V$ segue-se, pela Regra da Cadeia, $g'(b) \cdot f'(a) = \mathrm{id} \colon \mathbb{R}^m \to \mathbb{R}^m$ e $f'(a) \cdot g'(b) = \mathrm{id} \colon \mathbb{R}^n \to \mathbb{R}^n$. Assim, $g'(b) = f'(a)^{-1}$.

Em conseqüência do corolário acima, se $f\colon U \to V$ é um difeomorfismo entre os abertos $U \subset \mathbb{R}^m$ e $V \subset \mathbb{R}^n$ (bijeção, com f e f^{-1} ambas diferenciáveis), então, para todo $x \in U$, a derivada $f'(x)\colon \mathbb{R}^m \to \mathbb{R}^n$ é um isomorfismo. Em particular, $m = n$ e $U, V \subset \mathbb{R}^m$ são abertos do mesmo espaço euclidiano.

Existe um teorema de Topologia, devido a L.E.J. Brouwer ("teorema da invariância da dimensão") segundo o qual se $U \subset \mathbb{R}^m$ e $V \subset \mathbb{R}^n$ são abertos homeomorfos então $m = n$. A demonstração é bem mais difícil. (Veja [4], pág. 359.)

Um exemplo de homeomorfismo C^∞ cujo inverso não é diferenciável (no ponto 0) é dado pela função $f\colon \mathbb{R} \to \mathbb{R}$, onde $f(x) = x^3$. Isto mostra que um difeomorfismo não é a mesma coisa que um homeomorfismo diferenciável.

Ainda nesta ordem de idéias, é interessante observar que se $f\colon U \to V$ é um difeomorfismo entre abertos $U, V \subset \mathbb{R}^m$, sabendo-se que $f \in C^k$ $(1 \le k \le \infty)$ conclui-se que seu inverso $g = f^{-1}\colon V \to U$ (por definição, suposto apenas diferenciável) é também de classe C^k. Com efeito, se indicarmos com $\mathrm{GL}(\mathbb{R}^m)$ o conjunto das transformações lineares invertíveis de $\mathbb{R}^m$ em si mesmo e com $\mathrm{Inv}\colon \mathrm{GL}(\mathbb{R}^m) \to \mathcal{L}(\mathbb{R}^m; \mathbb{R}^m)$ a inversão de transformações lineares, a fórmula $g'(y) = [f'(g(y))]^{-1}$, estabelecida no Corolário 3 (onde $y \in V$), mostra que a aplicação derivada $g'\colon V \to \mathcal{L}(\mathbb{R}^m; \mathbb{R}^m)$ pode ser escrita como $g' = (\mathrm{Inv}) \circ f' \circ g$:

Como $\mathrm{Inv} \in C^\infty$, por um argumento já utilizado mais de uma vez, concluimos de $f \in C^k$ que $g \in C^k$. Podemos então enunciar o

Corolário 4. *Seja $f\colon U \to V$ uma bijeção de classe C^k $(k \ge 1)$ entre abertos $U, V \subset \mathbb{R}^m$. Se sua inversa $f^{-1}\colon V \to U$ é diferenciável então $f^{-1} \in C^k$. Diz-se então que f é um difeomorfismo de classe C^k.*

Quando $f\colon U \to \mathbb{R}^m$ é diferenciável no aberto $U \subset \mathbb{R}^m$ tem sentido, em cada ponto $x \in U$, considerar o determinante $\det Jf(x)$, da matriz jacobiana $Jf(x)$, chamado *o determinante jacobiano* de f no ponto x. Para que f seja um difeomorfismo é necessário que se tenha $\det Jf(x) \neq 0$ para todo $x \in U$. O Teorema da Aplicação Inversa, a ser demonstrado no parágrafo 8, mais adiante, fornece uma recíproca local para este fato.

Corolário 5 (Regras de derivação.) *Sejam $f, g\colon U \to \mathbb{R}^n$ diferenciáveis no ponto $a \in U \subset \mathbb{R}^m$ e c um número real. Então $f + g\colon U \to \mathbb{R}^n$ e $cf\colon U \to \mathbb{R}^n$ são diferenciáveis no ponto a, com $(f+g)'(a) = f'(a) + g'(a)$ e $(cf)'(a) = c \cdot f'(a)$. Quando $n = 1$ e $g(x) \neq 0$ para todo $x \in U$ então*

$f/g\colon U \to \mathbb{R}$ é diferenciável no ponto a, com $(f/g)'(a) = [g(a) \cdot f'(a) - f(a) \cdot g'(a)]/g(a)^2$. Finalmente, se $\varphi\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^p$ é bilinear, então $\varphi(f,g)\colon U \to \mathbb{R}^p$, definida por $x \mapsto \varphi(f(x), g(x))$, é diferenciável no ponto a com $[\varphi(f,g)]'(a)\colon v \mapsto \varphi(f'(a) \cdot v, g(a)) + \varphi(f(a), g'(a) \cdot v)$. Se $f, g \in C^k$ então $f+g$, $c \cdot f$, f/g, $\varphi(f,g) \in C^k$.

Em particular, quando $n = 1$ e $\varphi\colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ é a multiplicação de números reais então $\varphi(f,g) = f \cdot g$ e $(f \cdot g)'(a) = f'(a) \cdot g(a) + f(a) \cdot g'(a)$. Mais geralmente, como toda aplicação bilinear pode ser olhada como uma espécie de multiplicação, é razoável escrever $f(x) \cdot g(x)$ em vez de $\varphi(f(x), g(x))$ sempre que não houver perigo de mal entendidos. Então a regra de derivação acima fica como a "derivada de um produto":

$$(f \cdot g)'(a) = f'(a) \cdot g(a) + f(a) \cdot g'(a).$$

A interpretação desta fórmula é a seguinte: para todo vetor $v \in \mathbb{R}^m$, tem-se $(f \cdot g)'(a) \cdot v = (f'(a) \cdot v) \cdot g(a) + f(a) \cdot (g'(a) \cdot v)$.

Os três primeiros casos do corolário acima resultam imediatamente do Teorema 2 do Capítulo III. Tratemos explicitamente o caso de $\varphi(f,g) = \varphi \circ (f,g)$. Pela Regra da Cadeia, juntamente com o Exemplo 3 (no §1), temos, para todo $v \in \mathbb{R}^m$:

$$\begin{aligned}[\varphi(f,g)]'(a) \cdot v &= [\varphi \circ (f,g)]'(a) \cdot v = \\ &= \varphi'(f(a), g(a)) \cdot (f'(a) \cdot v, g'(a) \cdot v) = \\ &= \varphi(f'(a) \cdot v, g(a)) + \varphi(f(a), g'(a) \cdot v).\end{aligned}$$

Note-se ainda que $f + g = \alpha \circ (f,g)$, $c \cdot f = c^* \circ f$, $f/g = q \circ (f,g)$ e $\varphi(f,g) = \varphi \circ (f,g)$, onde $(f,g)\colon U \to \mathbb{R}^n \times \mathbb{R}^n$, $\alpha\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$, $c^*\colon \mathbb{R}^n \to \mathbb{R}^n$ e $q\colon \mathbb{R} \times (\mathbb{R} - \{0\}) \to \mathbb{R}$ são dadas, respectivamente, por $(f,g)(x) = (f(x), g(x))$, $\alpha(y,z) = y+z$, $c^*(y) = c \cdot y$ e $q(y,z) = y/z$. As aplicações α e c^* são lineares, logo de classe C^∞. A função q é também de classe C^∞ pois $q = m \circ (\mathrm{id}, \mathrm{inv})$ onde $\mathrm{id}\colon \mathbb{R} \to \mathbb{R}$ é a identidade, $\mathrm{inv}\colon \mathbb{R} - \{0\} \to \mathbb{R} - \{0\}$ é a inversão de números reais não nulos (matrizes invertíveis 1×1) e $m\colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ é a multiplicação de reais. Como m, id e inv são C^∞, q também o é. Segue-se então que $f, g \in C^k \Rightarrow f+g \in C^k$, $cf \in C^k$, $f/g \in C^k$ e $\varphi(f,g) \in C^k$.

A Regra da Cadeia é um instrumento útil para o cálculo das derivadas de aplicações diferenciáveis, e indispensável para a demonstração de muitas proposições de natureza teórica. Sua utilização pode, entretanto, ser dificultada pelo hábito arraigado de considerar derivadas como

números, em vez de transformações lineares, hábito proveniente do caso unidimensional, pois toda transformação linear $T\colon \mathbb{R} \to \mathbb{R}$ consiste na multiplicação por um número real fixo.

Exemplo 7. Sejam $f, g\colon U \to \mathbb{R}^n$ diferenciáveis (respect. de classe C^k) no aberto $U \subset \mathbb{R}^m$. Então $\xi\colon U \to \mathbb{R}$, definida por $\xi(x) = \langle f(x), g(x)\rangle$, é diferenciável (respect. de classe C^k), com $\xi'(x)\cdot v = \langle f'(x)\cdot v, g(x)\rangle + \langle f(x), g'(x)\cdot v\rangle$, para quaisquer $x \in U$, $v \in \mathbb{R}^m$. Em particular, tomando $f = g$, obtemos $\zeta(x) = |f(x)|^2 \Rightarrow \zeta'(x)\cdot v = 2\langle f'(x)\cdot v, f(x)\rangle$. Segue-se novamente da Regra da Cadeia que, em cada ponto $x \in U$ onde se tenha $f(x) \neq 0$, a função $\psi\colon U \to \mathbb{R}$, definida por $\psi(x) = |f(x)|$, isto é, $\psi(x) = \sqrt{\langle f(x), f(x)\rangle}$, é diferenciável, com $\psi'(x)\cdot v = \dfrac{\langle f'(x)\cdot v, f(x)\rangle}{|f(x)|}\cdot$

4 A fórmula de Taylor

No caso de uma aplicação $f\colon U \to \mathbb{R}^n$, definida no aberto $U \subset \mathbb{R}^m$, com $a, a+v \in U$, a fórmula de Taylor se escreve

$$(*)\qquad f(a+v) = f(a) + f'(a)\cdot v + \frac{1}{2} f''(a)\cdot v^2 + \cdots + \\ + \frac{1}{p!} f^{(p)}(a)\cdot v^p + r_p(v),$$

onde

$$f''(a)\cdot v^2 = \frac{\partial^2 f}{\partial v^2}(a) = \frac{\partial}{\partial v}\left(\frac{\partial f}{\partial v}\right)(a), \ldots,$$
$$f^p(a)\cdot v^p = \frac{\partial^p f}{\partial v^p} = \frac{\partial}{\partial v}\left(\frac{\partial^{p-1} f}{\partial v^{p-1}}\right)(a).$$

Como de costume, consideraremos três casos:

1) **Fórmula de Taylor Infinitesimal**. *Se f é p vezes diferenciável no ponto a então*

$$\lim_{v\to 0} \frac{r_p(v)}{|v|^p} = 0.$$

2) **Fórmula de Taylor com Resto de Lagrange**. *Sejam $[a, a+v] \subset U$, f de classe C^p, $p+1$ vezes diferenciável em cada ponto do*

segmento aberto $(a, a+v)$, *com* $|f^{(p+1)}(x)\cdot w^{p+1}| \le M\cdot|w|^{p+1}$ *para todo* $x \in (a, a+v)$ *e todo* $w \in \mathbb{R}^m$. *Então*

$$|r_p(v)| \le \frac{M}{(p+1)!}\,|v|^{p+1}.$$

3) **Fórmula de Taylor com Resto Integral**. *Se* f *é de classe* C^{p+1} *e* $[a, a+v] \subset U$, *então*

$$r_p(v) = \frac{1}{p!}\int_0^1 (1-t)^p\, f^{(p+1)}(a+tv)\cdot v^{p+1}\, dt.$$

Para cada $j = 0, 1, \ldots, p$, $f^{(j)}(a)\cdot v^j$ é o vetor de $\mathbb{R}^n$ cujas coordenadas são os números $d^{(j)} f_i(a)\cdot v^j$, onde $f_1, \ldots, f_n$ são as funções-coordenada de f. Logo a fórmula de Taylor (*) equivale a n igualdades numéricas que correspondem à fórmula de Taylor para funções reais, vista no §8 do Capítulo III. Como a integral que aparece acima é um vetor cujas n coordenadas são integrais numéricas semelhantes, apenas com $d^{(p+1)} f_i$ em vez de $f^{(p+1)}$, segue-se que as fórmulas 1) e 3) são conseqüências imediatas das fórmulas análogas para funções reais, já provadas no §8 do Capítulo III. Finalmente, a fórmula de Taylor com resto de Lagrange decorre do resultado análogo, provado para caminhos no §3 do Capítulo II, da maneira seguinte. Definindo o caminho $\varphi\colon [0,1] \to \mathbb{R}^n$ pondo $\varphi(t) = f(a+tv)$, vemos que φ é de classe C^p, $p+1$ vezes diferenciável no intervalo aberto $(0,1)$, com $\varphi'(t) = f'(a+tv)\cdot v$, $\varphi''(t) = f''(a+tv)\cdot v^2, \ldots, \varphi^{(p+1)}(t) = f^{(p+1)}(a+tv)\cdot v^{p+1}$, com $|\varphi^{(p+1)}(t)| \le M\cdot|v|^{p+1}$ para todo $t \in (0,1)$, em virtude da hipótese feita sobre f. A fórmula mencionada para caminhos diz então que

$$\varphi(1) = g(0) + \varphi'(0) + \frac{1}{2}\,\varphi''(0) + \cdots + \frac{1}{p!}\,\varphi^{(p)}(0) + r_p\,,$$

com

$$|r_p| \le M|v|^{p+1}/(p+1)!$$

Efetuando as devidas substituições, temos a fórmula de Taylor com resto de Lagrange, conforme enunciada em 2) acima.

Observação: Tem-se a *unicidade da fórmula de Taylor*, como foi provado no §8 do Capítulo III: se $f\colon U \to \mathbb{R}^n$ é p vezes diferenciável no

ponto $a \in \mathbb{R}^m$ e, para cada $i = 1, 2, \ldots, p$, é dada uma aplicação i-linear $\varphi_i \colon \mathbb{R}^m \times \cdots \times \mathbb{R}^m \to \mathbb{R}^n$, de tal modo que

$$f(a+v) = f(a) + \sum_{i=1}^{p} \frac{1}{i!} \varphi_i \cdot v^i + r_p(v),$$

com $\lim_{v \to 0} r_p(v)/|v|^p = 0$, então $\varphi_i \cdot v^i = f^{(i)}(a) \cdot v^i$ para todo $i = 1, \ldots, p$ e todo $v \in \mathbb{R}^m$.

5 A desigualdade do valor médio

Como as aplicações $f \colon U \to \mathbb{R}^n$ incluem, em particular, os caminhos, segue-se do Capítulo II, §3, que, para $n > 1$, não há um Teorema do Valor Médio sob forma de igualdade. Vale porém a desigualdade abaixo, caso particular da Fórmula de Taylor com resto de Lagrange, que estudaremos separadamente, dada sua importância.

Desigualdade do Valor Médio. *Dado $U \subset \mathbb{R}^m$ aberto, seja $f \colon U \to \mathbb{R}^n$ diferenciável em cada ponto do segmento de reta aberto $(a, a+v)$ e tal que sua restrição ao segmento fechado $[a, a+v] \subset U$ seja contínua. Se $|f'(x)| \leq M$ para todo $x \in (a, a+v)$ então $|f(a+v) - f(a)| \leq M \cdot |v|$.*

Demonstração: O caminho $\lambda \colon [0,1] \to \mathbb{R}^n$, definido por $\lambda(t) = f(a + tv)$, é (contínuo e) diferenciável no intervalo aberto $(0,1)$, com $\lambda(0) = f(a)$, $\lambda(1) = f(a+v)$ e, pela Regra da Cadeia, $\lambda'(t) = f'(a+tv) \cdot v$, donde $|\lambda'(t)| \leq |f'(a+tv)| \cdot |v| \leq M \cdot |v|$ para todo $t \in (0,1)$. Segue-se do Teorema do Valor Médio para caminhos (§3 do Capítulo II) que $|\lambda(1) - \lambda(0)| \leq M \cdot |v|$, ou seja, $|f(a+v) - f(a)| \leq M \cdot |v|$.

Corolário 1. *Seja $U \subset \mathbb{R}^m$ aberto e convexo. Se $f \colon U \to \mathbb{R}^n$ é diferenciável, com $|f'(x)| \leq M$ para todo $x \in U$ então f é Lipschitziana, com $|f(y) - f(x)| \leq M \cdot |y - x|$ para quaisquer $x, y \in U$.*

Com efeito, f é então contínua em U e $x, y \in U \Rightarrow [x, y] \subset U$, pela convexidade. Logo, vale a Desigualdade do Valor Médio.

Como vimos no Capítulo III (Exemplo 10), a convexidade de U é essencial para a validez do Corolário 1.

Corolário 2. *Se $f \colon U \to \mathbb{R}^n$ é diferenciável no aberto conexo $U \subset \mathbb{R}^n$ e $f'(x) = 0$ para todo $x \in U$ então f é constante.*

Fixemos $a \in U$ e ponhamos $A = \{x \in U; f(x) = f(a)\}$, $B = \{x \in U; f(x) \neq f(a)\}$. Como f é contínua, B é aberto. Afirmamos que A

também é aberto. Com efeito, dado $x \in A$, seja $\delta > 0$ o raio de uma bola de centro x, contida em U. Então $|v| < \delta \Rightarrow [x, x+v] \subset U \Rightarrow$ (pela Desigualdade do Valor Médio, com $f' \equiv 0$), $|f(x+v) - f(x)| = 0$, donde $f(x+v) = f(x) = f(a)$ e $x + v \in U$. Assim, $U = A \cup B$ é uma cisão. Como U é conexo e A é não-vazio, pois $a \in A$, vemos que $A = U$, isto é, $f(x) = f(a)$ para todo $x \in U$.

O Corolário abaixo fornece, quando $T = f'(a)$, uma estimativa para o resto $r(v) = f(a+v) - f(a) - f'(a) \cdot v$. Ele representa uma forma mais refinada da Desigualdade do Valor Médio, à qual se reduz tomando-se $T = 0$.

Corolário 3. *Sejam $U \subset \mathbb{R}^m$ aberto, $[a, a+v] \subset U$, $f\colon U \to \mathbb{R}^n$ diferenciável em cada ponto do intervalo aberto $(a, a+v)$, com $f|[a, a+v]$ contínua. Seja ainda $T\colon \mathbb{R}^m \to \mathbb{R}^n$ uma transformação linear. Se $|f'(x) - T| \leq M$ para todo $x \in (a, a+v)$ então*

$$|f(a+v) - f(a) - T \cdot v| \leq M \cdot |v|.$$

O Corolário acima resulta imediatamente de aplicar-se a Desigualdade do Valor Médio a $g\colon U \to \mathbb{R}^n$, definida por $g(x) = f(x) - T \cdot x$. Com efeito, tem-se, neste caso $g(a+v) - g(a) = f(a+v) - f(a) - T \cdot v$ e $g'(x) = f'(x) - T$.

Usaremos, em seguida, o Corolário 3 para provar que as aplicações de classe C^1 são uniformemente diferenciáveis nos compactos. Começaremos revendo a definição da aplicação diferenciável num conjunto, a fim de tornar mais explícita a notação para o resto $r(v) = f(a+v) - f(a) - f'(a) \cdot v$, deixando claro que ele depende não somente do acréscimo v como também do ponto a em questão.

Assim, uma aplicação $f\colon U \to \mathbb{R}^n$ se diz diferenciável no aberto $U \subset \mathbb{R}^m$ quando, para cada ponto $x \in U$, existe uma aplicação linear $f'(x)\colon \mathbb{R}^m \to \mathbb{R}^n$ tal que

$$f(x+v) - f(x) = f'(x) \cdot v + r_x(v), \text{ onde } \lim_{v \to 0} \frac{r_x(v)}{|v|} = 0.$$

O limite acima significa que, dado $\varepsilon > 0$, se pode obter, para cada $x \in U$, um número $\delta = \delta(x, \varepsilon)$, *dependendo de x e de ε*, tal que $|v| < \delta \Rightarrow |r_x(v)| < \varepsilon|v|$.

Diz-se que a aplicação diferenciável $f\colon U \to \mathbb{R}^n$ é *uniformemente diferenciável* num subconjunto $X \subset U$ quando, para todo $\varepsilon > 0$ dado,

existe $\delta > 0$ tal que $|v| < \delta \Rightarrow |f(x+v) - f(x) - f'(x) \cdot v| < \varepsilon|v|$, *seja qual for* $x \in X$ (com $x + v \in U$).

Corolário 4. *Uma aplicação* $f \colon U \to \mathbb{R}^n$, *de classe* C^1, *é uniformemente diferenciável em qualquer compacto* $K \subset U$.

Existe $\delta > 0$ tal que $x \in K,\ \ |v| < \delta \Rightarrow [x, x+v] \subset U$. (Corolário 2 do Teorema 26, Capítulo I.) Dado arbitrariamente $\varepsilon > 0$ podemos reduzir δ, se for necessário, de modo que $x \in K,\ \ |v| < \delta \Rightarrow |f'(x+v) - f'(x)| < \varepsilon$, já que $f' \colon U \to \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$ é contínua. (Teorema 21a do Capítulo I.) Daí se segue, obviamente, que $x \in K,\ \ |v| < \delta,\ \ 0 \leq t \leq 1 \Rightarrow |f'(x+tv) - f'(x)| < \varepsilon$, ou seja, $y \in [x, x+v] \Rightarrow |f'(y) - f'(x)| < \varepsilon$. Pelo Corolário 3 acima, usado com $T = f'(x)$, concluimos que $|f(x+v) - f(x) - f'(x) \cdot v| \leq \varepsilon|v|$ sempre que $|v| < \delta$, seja qual for $x \in K$, o que prova o Corolário 4.

Corolário 5. *Sejam* $U \subset \mathbb{R}^m$ *aberto e* $c \in U$. *Se a aplicação contínua* $f \colon U \to \mathbb{R}^n$ *é diferenciável em* $U - \{c\}$ *e existe* $\lim_{x \to c} f'(x) = T \in \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$, *então* f *é diferenciável no ponto* c, *com* $f'(c) = T$.

Em virtude da definição de limite, dado $\varepsilon > 0$, existe $\delta > 0$ tal que $0 < |v| < \delta \Rightarrow |f'(c+tv) - T| < \varepsilon$ seja qual for $t \in (0, 1)$. Podemos supor δ tão pequeno que $|v| < \delta \Rightarrow [c, c+v] \subset U$. (Basta tomar $\delta =$ raio de uma bola de centro c, contida em U.) Então, pelo Corolário 3 acima, pondo $r(v) = f(c+v) - f(c) - T \cdot v$, temos $|r(v)| \leq \varepsilon|v|$ sempre que $0 < |v| < \delta$. Isto mostra que f é diferenciável no ponto c, com $f'(c) = T$.

Observação: O Corolário 5 mostra a vantagem de ter demonstrado o Corolário 3 em toda a sua generalidade. [Uma transformação linear arbitrária T, em vez de $f'(a)$, e f suposta apenas diferenciável no segmento aberto $(a, a+v)$.] Um exemplo de como o Corolário 5 pode ser utilizado, para funções de uma variável, é visto no Volume 1, página 275. (Exemplo 20.)

6 Seqüências de aplicações diferenciáveis

Provaremos neste parágrafo o teorema clássico, segundo o qual o limite uniforme de uma seqüência de aplicações diferenciáveis é ainda uma aplicação diferenciável desde que a seqüência das derivadas também convirja uniformemente. Como no caso de uma variável, a Desigualdade do Valor Médio é o instrumento mais importante da demonstração.

As definições básicas sobre convergência uniforme e os teoremas demonstrados no §1 e 2 do Capítulo 10, Volume 1, para funções de uma variável, se aplicam, com modificações evidentes, ao caso $f\colon U \to \mathbb{R}^n$, $U \subset \mathbb{R}^m$. Uma exceção notável é o teorema sobre a derivada do limite de uma seqüência de funções (Teorema 7 daquele capítulo), pois em sua demonstração foi utilizado o Teorema do Valor Médio sob forma de igualdade, que não vale para $n > 1$, e cuja versão em dimensões superiores envolve uma hipótese de convexidade, que é automática quando o domínio é um intervalo. Começaremos aqui reexpondo o conceito de convergência uniforme.

Diz-se que uma seqüência de aplicações $f_k\colon X \to \mathbb{R}^n$, definidas num conjunto X, *converge uniformemente* para uma aplicação $f\colon X \to \mathbb{R}^n$ quando, para qualquer $\varepsilon > 0$ dado, pode-se obter $k_0 \in \mathbb{N}$ tal que $k > k_0 \Rightarrow |f_k(x) - f(x)| < \varepsilon$, seja qual for $x \in X$.

Resulta imediatamente da definição que a afirmação "$\lim f_k = f$ uniformemente em X"não depende da norma que se considera no espaço euclidiano. Tem-se $f_k \to g$ uniformemente em X se, e somente se, para cada $i = i, \ldots, n$, tem-se $\lim\limits_{k\to\infty} f_{ki} = g_i$ uniformemente em X, onde $f_{k1}, \ldots, f_{kn}\colon X \to \mathbb{R}$ são as funções-coordenada da aplicação $f_k\colon X \to \mathbb{R}^n$ e $g_1, \ldots, g_n\colon X \to \mathbb{R}$ são as funções-coordenada de g.

Como o espaço $\mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$ possui uma norma, a definição de convergência uniforme faz sentido também para uma seqüência de aplicações $g_k\colon X \to \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$. Neste caso, tem-se $g_k \to g$ uniformemente em X se, e somente se, para todo $\varepsilon > 0$ dado, existe $k_0 \in \mathbb{N}$ tal que $k > k_0 \Rightarrow |g_k(x) \cdot v - g(x) \cdot v| \leq \varepsilon \cdot |v|$ para todo $x \in X$ e todo vetor $v \in \mathbb{R}^m$. Com efeito, pela definição de norma da transformação linear $g_k(x) - g(x)\colon \mathbb{R}^m \to \mathbb{R}^n$, tem-se $|g_k(x) - g(x)| \leq \varepsilon \Leftrightarrow |g_k(x) \cdot v - g(x) \cdot v| \leq \varepsilon |v|$ para todo $v \in \mathbb{R}^m$.

Seja agora $X \subset \mathbb{R}^m$. Diz-se que a seqüência de aplicações $f_k\colon X \to \mathbb{R}^n$ converge de modo *localmente uniforme* em X para uma aplicação $f\colon X \to \mathbb{R}^n$ quando cada ponto $x \in X$ é centro de uma bola aberta B tal que $f_k \to f$ uniformemente em $B \cap X$. Isto equivale a dizer que X está contido numa reunião de abertos $U \subset \mathbb{R}^m$ tais que em cada $U \cap X$ a convergência $f_k \to f$ é uniforme.

Evidentemente, convergência uniforme $\Rightarrow$ convergência localmente uniforme $\Rightarrow$ convergência simples (isto é, $\lim\limits_{k\to\infty} f_k(x) = f(x)$ para todo $x \in X$). As implicações contrárias são falsas.

Exemplo 8. Para cada $k \in \mathbb{N}$, seja $f_k \colon \mathbb{R} \to \mathbb{R}$ dada por $f_k(x) = x/k$. A seqüência de funções f_k converge de modo localmente uniforme em $\mathbb{R}$ para a função identicamente nula, mas não converge uniformemente. (Veja Volume I, páginas 362 e 366.)

Critério de Cauchy. *A fim de que uma seqüência de aplicações $f_k \colon X \to \mathbb{R}^n$ convirja uniformemente em X é necessário e suficiente que, para todo $\varepsilon > 0$ dado, se possa obter $k_0 \in \mathbb{N}$ tal que $j, k > k_0 \Rightarrow |f_j(x) - f_k(x)| < \varepsilon$, seja qual for $x \in X$.*

A demonstração dada na página 368 do Volume I se aplica, palavra por palavra. Como conseqüência, temos o

Teste de Weierstrass. *Se, para cada $k \in \mathbb{N}$ e cada $x \in X$, tem-se $|f_k(x)| \leq c_k$, onde $\Sigma\, c_k$ é uma série convergente de números reais positivos, então a série $\Sigma\, f_k$, cujos termos são aplicações $f_k \colon X \to \mathbb{R}^n$, converge uniformemente em X.*

No caso acima, a série $\Sigma\, f_k$ converge absoluta e uniformemente, isto é, a série $\Sigma |f_k(x)|$ também converge uniformemente em X, já que cumpre a mesma condição que $\Sigma\, f_k$ no teste de Weierstrass.

Como no caso de uma variável, a proposição abaixo é uma das razões da importância do conceito de convergência uniforme.

Teorema da Continuidade do Limite Uniforme. *Se a seqüência de aplicações $f_k \colon X \to \mathbb{R}^n$, definidas no conjunto $X \subset \mathbb{R}^m$ e contínuas no ponto $a \in X$, converge uniformemente em X para a aplicação $f \colon X \to \mathbb{R}^n$, então f é contínua no ponto a.*

Demonstração: Dado $\varepsilon > 0$, existe $k_0 \in \mathbb{N}$ tal que $k > k_0 \Rightarrow |f_k(x) - f(x)| < \varepsilon/3$, seja qual for $x \in X$. Fixemos um índice $k > k_0$. Existe $\delta > 0$ tal que $|x - a| < \delta, \quad x \in X \Rightarrow |f_k(x) - f_k(a)| < \varepsilon/3$, pela continuidade de f_k no ponto a. Segue-se então que $x \in X, \quad |x - a| < \delta \Rightarrow |f(x) - f(a)| \leq |f(x) - f_k(x)| + |f_k(x) - f_k(a)| + |f_k(a) - f(a)| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon$. Logo, f é contínua no ponto a.

O teorema que temos em vista, sobre a diferenciabilidade do limite de uma seqüência de aplicações diferenciáveis, se apóia no lema abaixo, do qual é uma extensão.

Lema. *Seja $U \subset \mathbb{R}^m$ um aberto convexo e limitado. Se a seqüência de aplicações diferenciáveis $f_k \colon U \to \mathbb{R}^n$ converge num ponto $c \in U$ e a seqüência das derivadas $f_k' \colon U \to \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$ converge uniformemente em U para uma aplicação $g \colon U \to \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$, então (f_k) con-*

verge uniformemente em U para uma aplicação $f: U \to \mathbb{R}^n$, a qual é diferenciável, com $f' = g$.

Demonstração: Em primeiro lugar, dado $\varepsilon > 0$, existe $k_0 \in \mathbb{N}$ tal que

$$(1) \qquad j, k > k_0 \Rightarrow |f_j'(x) - f_k'(x)| < \varepsilon \text{ para todo } x \in U.$$

Como U é convexo, resulta do Corolário 1 da Desigualdade do Valor Médio, aplicado a $f_j - f_k$, que, para quaisquer $x, y \in U$:

$$(2) \qquad j, k > k_0 \Rightarrow |f_j(y) - f_k(y) - [f_j(x) - f_k(x)]| \leq \varepsilon |y - x|.$$

Tomando $x = c$, vemos que, para todo $y \in U$:

$$j, k > k_0 \Rightarrow |f_j(y) - f_k(y)| \leq |f_j(c) - f_k(c)| + \varepsilon |y - c|.$$

Usando o Critério de Cauchy, o fato de U ser limitado e a convergência de $(f_k(c))$, concluimos daí que (f_k) converge uniformemente para uma aplicação $f: U \to \mathbb{R}^n$. Mostremos agora que f é diferenciável em todo ponto $x \in U$, com $f'(x) = g(x)$. Com efeito, fazendo $j \to \infty$ em (2) e escrevendo $y = x + v$, obtemos

$$(3) \qquad k > k_0 \Rightarrow |f(x+v) - f(x) - [f_k(x+v) - f_k(x)]| \leq \varepsilon \cdot |v|.$$

Como cada f_k é diferenciável no ponto x, para todo $k \in \mathbb{N}$ existe $\delta_k(x) > 0$ tal que

$$(4) \qquad |v| < \delta_k(x) \Rightarrow |f_k(x+v) - f_k(x) - f_k'(x) \cdot v| < \varepsilon |v|.$$

Fazendo $j \to \infty$ em (1) vem:

$$(5) \qquad k > k_0 \Rightarrow |f_k'(x) - g(x)| \leq \varepsilon \text{ para todo } x \in U.$$

Para mostrar que $g(x)$ é a derivada de f no ponto x, dado $\varepsilon > 0$ tomamos k_0 como em (1), escolhemos algum inteiro $k > k_0$ e pomos $\delta = \delta_k(x)$. Então, em virtude de (3), (4) e (5),

$$|v| < \delta \Rightarrow |f(x+v) - f(x) - g(x) \cdot v| \leq$$
$$\leq |f(x+v) - f(x) - [f_k(x+v) - f_k(x)]| +$$
$$+ |f_k(x+v) - f_k(x) - f_k'(x) \cdot v| + |f_k'(x) \cdot v - g(x) \cdot v| < 3\varepsilon \cdot |v|,$$

o que prova o lema.

Teorema da Derivação Termo a Termo. *Seja $U \subset \mathbb{R}^m$ um aberto conexo. Se a seqüência de aplicações diferenciáveis $f_k\colon U \to \mathbb{R}^n$ converge num ponto $c \in U$ e a seqüência das derivadas $f_k'\colon U \to \mathcal{L}(\mathbb{R}^m;\mathbb{R}^n)$ converge de modo localmente uniforme para uma aplicação $g\colon U \to \mathcal{L}(\mathbb{R}^m;\mathbb{R}^n)$, então (f_k) converge de modo localmente uniforme para uma aplicação $f\colon U \to \mathbb{R}^n$, a qual é diferenciável, com $f' = g$.*

Concisamente: *vale $D(\lim f_k) = \lim Df_k$, desde que a convergência das derivadas seja localmente uniforme.*

Demonstração: Podemos escrever $U = \cup B_\alpha$, onde B_α é uma bola aberta na qual as derivadas f_k' convergem uniformemente. Pelo Lema, se (f_k) converge em algum ponto de B_α então (f_k) converge uniformemente em B_α. Tem-se assim uma cisão $U = A \cup B$, onde A é a reunião das bolas B_α nas quais (f_k) converge uniformemente e B é a reunião das demais bolas, nas quais não há convergência em ponto algum. Como U é conexo e A não é vazio (pois se $c \in B_\alpha$ então $B_\alpha \subset A$), segue-se que $A = U$, logo (f_k) converge de modo localmente uniforme em U para uma aplicação $f\colon U \to \mathbb{R}^n$. Pelo Lema, tem-se $f' = g$.

Observação: Mesmo supondo que $f_k' \to g$ uniformemente no aberto conexo U, não se segue de $\lim f_k(c) = f(c)$, que $f_k \to f$ uniformemente em U, como se vê tomando as funções $f_k\colon \mathbb{R} \to \mathbb{R}$, $f_k(x) = x/k$. Neste caso, U é ilimitado. Mas a validez do Lema acima, no qual se conclui a convergência *uniforme* de f_k em U a partir da convergência uniforme das derivadas mais a convergência de $(f_k(c))$ para algum $c \in U$ não é assegurada supondo-se U conexo e limitado. O essencial é que exista um número real $M > 0$ tal que dois pontos quaisquer de U podem ser ligados por uma poligonal de comprimento $\leq M$, contida em U. Quando U é convexo e limitado, isto ocorre.

Corolário. (Derivação Termo a Termo para Séries.) *Seja $U \subset \mathbb{R}^m$ aberto e conexo. Se a série Σf_k, de aplicações diferenciáveis $f_k\colon U \to \mathbb{R}^n$, converge num ponto $c \in U$ e a série $\Sigma f_k'$ das derivadas converge de modo localmente uniformemente em U, para a soma $g = \Sigma f_k'$, então tem-se convergência localmente uniforme $f = \Sigma f_k$ em U, a soma $f\colon U \to \mathbb{R}^n$ é diferenciável e $f' = g$. Em suma: $D(\Sigma f_k) = \Sigma Df_k$ se a série das derivadas convergir de modo localmente uniforme.*

7 Aplicações fortemente diferenciáveis

Existe uma noção de diferenciabilidade, correspondendo ao que seria classe C^1, mas onde se supõe que a aplicação é diferenciável num único ponto. Trata-se da noção de diferenciabilidade forte, que discutiremos mais abaixo.

Começamos com uma proposição que generaliza ligeiramente o Teorema 4a do Capítulo III. Trata-se de um primeiro e simples exemplo de como atua a hipótese de diferenciabilidade de uma aplicação: a partir do comportamento de derivada $f'(a)$ se podem obter informações sobre o comportamento de $f(x)$ para x próximo de a.

Teorema 3. *Seja $f\colon U \to \mathbb{R}^n$ definida no aberto $U \subset \mathbb{R}^m$, diferenciável no ponto $a \in U$.*

1º) *Se a transformação linear $f'(a)\colon \mathbb{R}^m \to \mathbb{R}^n$ é injetiva então existem $\delta > 0$ e $c > 0$ tais que $|x-a| < \delta \Rightarrow |f(x) - f(a)| \geq c|x-a|$;*

2º) *Se $f'(a)\colon \mathbb{R}^m \to \mathbb{R}^n$ é sobrejetiva então em qualquer bola de centro a existem pontos x tais que $|f(a)| < |f(x)|$ e (a menos que seja $f(a) = 0$) pontos y tais que $|f(y)| < |f(a)|$.*

Demonstração: 1º) Podemos argumentar como no Teorema 4a do Capítulo III, ou no Lema contido no Exemplo 4, acima. Observando que $f'(a)$ é um isomorfismo de $\mathbb{R}^m$ sobre o subespaço $E = f'(a) \cdot \mathbb{R}^m \subset \mathbb{R}^n$, pomos $2c = 1/|f'(a)^{-1}|$, onde $f'(a)^{-1}\colon E \to \mathbb{R}^m$ é o inverso de $f'(a)$. Então $|f'(a) \cdot v| \geq 2c \cdot |v|$ para todo $v \in \mathbb{R}^m$. Existe, pela diferenciabilidade de f no ponto a, um número $\delta > 0$ tal que $x \in U$, $|x-a| < \delta \Rightarrow f(x) - f(a) = f'(a) \cdot (x-a) + r(x-a)$, com $|r(x-a)| < c|x-a|$, e daí $|f(x)-f(a)| \geq |f'(a)\cdot(x-a)| - |r(x-a)| \geq 2c\cdot|x-a| - c\cdot|x-a| = c\cdot|x-a|$.

2º) Se não existissem pontos x, arbitrariamente próximos de a, com $|f(a)| < |f(x)|$ então $a \in U$ seria um ponto de máximo local para a função $\zeta\colon U \to \mathbb{R}$, $\zeta(x) = |f(x)|$ (norma euclidiana). [Neste caso, deveria ser $f(a) \neq 0$ pois do contrário f seria zero em todos os pontos de uma bola de centro a, logo $f'(a) = 0$ não seria sobrejetiva.] Então ζ é diferenciável no ponto a (veja o Exemplo 7), com $0 = \zeta'(a) \cdot v = \langle f'(a) \cdot v, f(a)\rangle / |f(a)|$ para todo $v \in \mathbb{R}^m$. Daí resultaria, em particular, que $f'(a)\cdot v \neq f(a)$ para qualquer v e $f'(a)$ não seria sobrejetiva. Quanto à segunda afirmação deste item, se tivermos $f(a) \neq 0$ e não existirem pontos y arbitrariamente próximos de a com $|f(y)| < |f(a)|$ então a

será um mínimo local para a função $\zeta(x) = |f(x)|$, o que leva a uma contradição como acima.

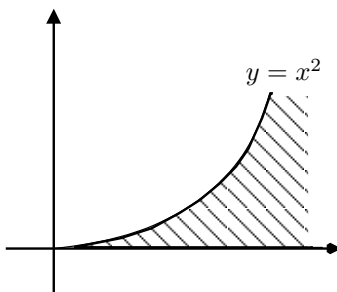
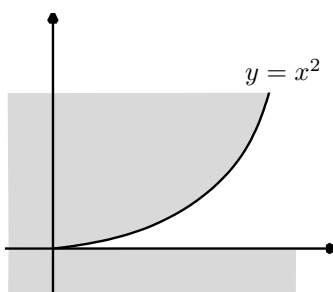
Corolário. *Se $f\colon U \to \mathbb{R}^n$ é diferenciável no ponto a e $f'(a)$ é injetiva então existe uma bola B de centro a tal que $x \in B,\quad x \neq a \Rightarrow f(x) \neq f(a)$.*

Em relação ao Teorema 3, cabe a pergunta: supondo $f'(a)$ injetiva, existe alguma bola B de centro a, tal que $|f(y) - f(x)| \geq c|y - x|$, sejam quais forem $x, y \in B$ ou, pelo menos, que a restrição $f|B$ seja injetiva? Cabe também a pergunta: supondo $f'(a)$ sobrejetiva pode-se concluir que $f(a)$ pertença ao interior da imagem $f(U)$?

Sem hipóteses adicionais, a resposta a estas perguntas é a mesma: não. Já sabíamos desde o Volume 1 (veja o Exemplo 12, página 266) que uma função $f\colon \mathbb{R} \to \mathbb{R}$ pode ser derivável, com $f'(a) \neq 0$ [isto é o que significa $f'(a)$ ser injetiva – ou sobrejetiva, tanto faz – em dimensão 1] sem que f seja injetiva em vizinhança alguma de a. Mas naquele exemplo (como em qualquer outro que se busque em dimensão 1), o ponto $f(a)$ pertence ao interior de $f(I)$, seja qual for o intervalo aberto I de centro a.

O exemplo abaixo exibe uma aplicação $f\colon \mathbb{R}^2 \to \mathbb{R}^2$, diferenciável na origem $0 \in \mathbb{R}^2$, com $f'(0) =$ identidade: $\mathbb{R}^2 \to \mathbb{R}^2$, mas f não é injetiva em vizinhança alguma de 0 nem tampouco $f(0)$ pertence ao interior de $f(U)$, seja qual for o aberto U contendo 0.

Exemplo 9. Definimos $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ pondo $f(x, y) = (x, y)$, salvo quando $x > 0$ e $0 < y < x^2$, em cujo caso poremos $f(x, y) = (x, x^2)$. A aplicação f é, admitidamente, descontínua nos pontos $(x, 0)$, com $x > 0$, mas é contínua nos demais pontos, inclusive na origem $0 = (0, 0)$, sendo até mesmo diferenciável ali, com $f'(0) =$ identidade. Para verificar esse fato, observemos que se pode escrever, para todo $v = (x, y) \in \mathbb{R}^2$, $f(v) - f(0) = v + r(v)$, onde $r(v) = (0, x^2 - y)$ se $x > 0$ e $0 < y < x^2$, e $r(v) = 0$ nos demais casos. Então a segunda coordenada de $r(v)$ está sempre compreendida entre 0 e x^2. Como a primeira é sempre zero, resulta que $\lim\limits_{v \to 0} r(v)/|v| = \lim r(v)/\sqrt{x^2 + y^2} = 0$. Em qualquer aberto contendo zero está contido um segmento de reta vertical de extremos $(x, 0)$ e (x, x^2), com $x > 0$, o qual é transformado por f num único ponto (x, x^2). Logo f não é injetiva em vizinhança alguma de 0. Além disso, como nenhum ponto (x, y), com $x > 0$ e $0 < y < x^2$, pertence à imagem de f, vemos que $f(0) = 0$ não é interior a nenhuma imagem

$f(U)$, onde $U \subset \mathbb{R}^2$ é um aberto contendo 0.

A parte hachurada transforma-se por f na curva $y = x^2$, $x > 0$.

A parte sombreada é a imagem de f.

Observação: Não seria difícil modificar um pouco o exemplo acima, de modo a obter uma aplicação *contínua* $f\colon \mathbb{R}^2 \to \mathbb{R}^2$, diferenciável na origem, com $f'(0) =$ identidade, tal que em nenhuma bola de centro 0 f é injetiva. Entretanto, mostraremos no exemplo 36, §10 do Capítulo VII, com auxílio da teoria do grau, que se $f\colon U \to \mathbb{R}^m$ é contínua no aberto $U \subset \mathbb{R}^m$ e possui, no ponto a e U, uma derivada $f'(a)\colon \mathbb{R}^m \to \mathbb{R}^m$ que é um isomorfismo, então $f(a) \in \operatorname{int} f(U)$. Isto mostra que a descontinuidade de f no Exemplo 9 é essencial para termos $f(a) \notin \operatorname{int}(U)$.

Uma aplicação $f\colon U \to \mathbb{R}^n$, definida no aberto $U \subset \mathbb{R}^m$, chama-se *fortemente diferenciável* no ponto $a \in U$ quando existe uma transformação linear $T\colon \mathbb{R}^m \to \mathbb{R}^n$ tal que, para $x, y \in U$ vale

$$f(x) - f(y) = T \cdot (x - y) + \rho_a(x, y)|x - y|, \text{ onde } \lim_{(x,y)\to(a,a)} \rho_a(x, y) = 0.$$

Tomando $y = a$, vemos que toda aplicação fortemente diferenciável no ponto a é diferenciável nesse ponto, com $T = f'(a)$. Assim a condição acima se torna:

$$f(x) - f(y) = f'(a) \cdot (x - y) + \rho_a(x, y)|x - y|$$

onde $\rho_a(x, y)$ cumpre a seguinte condição: dado qualquer $\varepsilon > 0$, pode-se achar $\delta > 0$ tal que $x, y \in B(a, \delta) \Rightarrow |\rho_a(x, y)| < \varepsilon$.

Quando $m = n = 1$, uma função $f\colon I \to \mathbb{R}$, definida no intervalo $I \subset \mathbb{R}$, é fortemente diferenciável no ponto $a \in I$ quando, dados $x \neq y$ em I, a secante ao gráfico de f que passa pelos pontos $(x, f(x))$ e $(y, f(y))$ tende para a tangente no ponto $(a, f(a))$ quando $x \to a$ e $y \to a$. Na

definição usual de derivada, o ponto a é mantido fixo: toma-se a secante ao gráfico passando pelos pontos $(a, f(a))$ e $(x, f(x))$ e faz-se $x \to a$.

Segue-se da definição que se $f\colon U \to \mathbb{R}^n$ é fortemente diferenciável no ponto a então, para todo $\varepsilon > 0$, pode-se obter $\delta > 0$ tal que, para $x, y \in B(a;\delta)$, f satisfaz à condição de Lipschitz

$$|f(x) - f(y)| \leq (|f'(a)| + \varepsilon)|x - y|.$$

Com efeito, dado $\varepsilon > 0$, existe $\delta > 0$ tal que

$$\begin{aligned} x, y \in B(a;\delta) \Rightarrow & |f(x) - f(y)| \leq |f'(a)| \cdot |x - y| + \varepsilon|x - y| = \\ & = (|f'(a)| + \varepsilon)|x - y|. \end{aligned}$$

Em particular, f é contínua em todos os pontos de uma bola de centro a.

Outra conseqüência da definição é o

Teorema 4. *Se $f\colon U \to \mathbb{R}^n$ é fortemente diferenciável no ponto a e $f'(a)\colon \mathbb{R}^m \to \mathbb{R}^n$ é injetiva então existem $\delta > 0$ e $c > 0$ tais que $x, y \in B(a;\delta) \Rightarrow |f(x) - f(y)| \geq c \cdot |x - y|$. Segue-se que f é um homeomorfismo da bola $B(a;\delta)$ sobre sua imagem em $\mathbb{R}^n$ e, em particular, que f é injetiva nessa bola.*

Demonstração: Sendo $f'(a)$ injetiva, sabemos que existe $c > 0$ tal que $|f'(a) \cdot v| \geq 2c|v|$ para todo $v \in \mathbb{R}^m$. Tomando $\varepsilon = c$, obtemos $\delta > 0$ tal que $x, y \in B(a;\delta) \Rightarrow f(x) - f(y) = f'(a) \cdot (x - y) + r_a(x, y)$, com $|r_a(x, y)| < c|x - y|$. Então

$$\begin{aligned} x, y \in B(a;\delta) \Rightarrow |f(x) - f(y)| & \geq |f'(a) \cdot (x - y)| - |r_a(x, y)| \geq \\ & \geq 2c \cdot |x - y| - c \cdot |x - y| = c|x - y|. \end{aligned}$$

Conseqüentemente, f aplica a bola $B(a;\delta)$ bijetivamente sobre sua imagem $Y = f(B(a;\delta))$ e a inversa $f^{-1}\colon Y \to B(a;\delta)$ cumpre a condição de Lipschitz $|f^{-1}(w) - f^{-1}(z)| \leq c^{-1}|w - z|$. Assim, f^{-1} é contínua, logo $f\colon B(a;\delta) \to Y$ é um homeomorfismo.

Observação: Provaremos, no §11, que se $f\colon U \to \mathbb{R}^n$ é fortemente diferenciável no ponto $a \in U$ e a derivada $f'(a)\colon \mathbb{R}^m \to \mathbb{R}^n$ é sobrejetiva então $f(a) \in \operatorname{int} f(U)$.

Sejam $f\colon U \to \mathbb{R}^n$, $a \in U \subset \mathbb{R}^m$, $T \in \mathcal{L}(\mathbb{R}^m;\mathbb{R}^n)$ e $x, y \in U$. Escrevendo

$$f(x) - f(y) = T \cdot (x - y) + r_a(x, y), f(x) - f(a) = T \cdot (x - a) + r_a(x)$$
$$\text{e} \quad f(y) - f(a) = T \cdot (y - a) + r_a(y),$$

obtemos, por subtração membro a membro das 2 últimas igualdades:

$$r_a(x, y) = r_a(x) - r_a(y).$$

Daí resulta que $|r_a(x, y)| \leq \varepsilon|x - y| \Leftrightarrow |r_a(x) - r_a(y)| \leq \varepsilon \cdot |x - y|$. Podemos então enunciar o

Teorema 5. *A aplicação $f : U \to \mathbb{R}^n$ é fortemente diferenciável no ponto $a \in U$ se, e somente se, é diferenciável nesse ponto e, para todo $\varepsilon > 0$, se pode achar $\delta > 0$ tal que o resto $r_a(x) = f(x) - f(a) - f'(a) \cdot (x - a)$ satisfaz a condição de Lipschitz $|r_a(x) - r_a(y)| \leq \varepsilon|x - y|$ para $x, y \in B(a; \delta)$.*

A proposição abaixo mostra que a única diferença entre a diferenciabilidade forte e a continuidade da derivada é que a primeira faz sentido mesmo quando a aplicação é diferenciável num único ponto.

Teorema 6. *Seja $f : U \to \mathbb{R}^n$ uma aplicação diferenciável. A fim de que f seja fortemente diferenciável no ponto $a \in U$, é necessário e suficiente que a aplicação derivada $f' : U \to \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$ seja contínua no ponto a.*

Demonstração: Suponhamos f' contínua no ponto a. Então a aplicação $r_a : U \to \mathbb{R}^n$, definida por $r_a(x) = f(x) - f(a) - f'(a) \cdot (x - a)$, é diferenciável, e sua derivada é contínua no ponto a. Como $r'_a(x) = f'(x) - f'(a)$, temos $r'_a(a) = 0$. Logo, para todo $\varepsilon > 0$, existe $\delta > 0$ tal que $B(a; \delta) \subset U$ e $x \in B(a; \delta) \Rightarrow |r'_a(x)| < \varepsilon$. Pelo Corolário 1 da Desigualdade do Valor Médio, $x, y \in B(a; \delta) \Rightarrow |r_a(x) - r_a(y)| \leq \varepsilon|x - y|$. Segue-se do Teorema 5 que f é fortemente diferenciável no ponto a.

Reciprocamente, seja f fortemente diferenciável no ponto a. Dado arbitrariamente $\varepsilon > 0$, devemos achar $\delta > 0$ tal que $|x - a| < \delta$ implique $|[f'(x) - f'(a)] \cdot u| < \varepsilon$ para todo vetor unitário $u \in \mathbb{R}^m$. Somando as igualdades

$$f(x) - f(y) = f'(a) \cdot (x - y) + r_a(x, y) \qquad \text{e}$$
$$f(y) - f(x) = f'(x) \cdot (y - x) + r_x(y),$$

obtemos:

$$(*) \qquad [f'(x) - f'(a)] \cdot (y - x) = -[r_a(x, y) + r_x(y)],$$

Dado $\varepsilon > 0$, existe $\delta > 0$ tal que

$$x, y \in B(a; 2\delta) \Rightarrow |r_a(x, y)| \leq \frac{\varepsilon}{2} |y - x|$$

e, para todo $x \in U$, existe $\eta_x > 0$ tal que $|y - x| \leq \eta_x$ implica $|r_x(y)| \leq (\varepsilon/2)|y - x|$.
Sem perda de generalidade, podemos supor $\eta_x < \delta$. Sejam então $x \in U$ tal que $|x - a| < \delta$ e $u \in \mathbb{R}^m$ um vetor unitário. Pondo $y = x + \eta_x u$, temos $|y - a| < 2\delta$ e $|y - x| < \eta_x$, donde

$$|r_a(x, y)| \leq \frac{\varepsilon}{2}\,|x - y| < \frac{\varepsilon}{2}\,\eta_x \text{ e } |r_x(y)| \leq \frac{\varepsilon}{2}\,|y - x| < \frac{\varepsilon}{2}\,\eta_x\,.$$

Segue-se de (*) que $|[f'(x) - f'(a)] \cdot u| < \varepsilon$, como queríamos demonstrar.

Exemplo 10. Para obter exemplo de uma função fortemente diferenciável no ponto $a \in \mathbb{R}$, mas não diferenciável em vizinhança alguma de a, considere uma função $f\colon \mathbb{R} \to \mathbb{R}$, de classe C^1, cujo gráfico não contenha três pontos colineares. Considere ainda uma seqüência decrescente $a_n \to a$ e defina uma nova função $g\colon \mathbb{R} \to \mathbb{R}$, que coincide com f nos pontos a_n e fora do intervalo (a, a_1). Em cada intervalo $[a_{n+1}, a_n]$, faça o gráfico de g ser retilíneo.

8 O teorema da aplicação inversa

Sejam $U, V \subset \mathbb{R}^m$ abertos. Um difeomorfismo $f\colon U \to V$ é, como sabemos, uma bijeção diferenciável cuja inversa é diferenciável. Em particular, f é um homeomorfismo entre U e V. O exemplo clássico de $f\colon \mathbb{R} \to \mathbb{R}$, $f(x) = x^3$, mostra que um homeomorfismo pode ser diferenciável sem que seu inverso o seja. Quando o difeomorfismo f é de classe C^k, seu inverso também o é, pelo Corolário 4 da Regra da Cadeia. Evidentemente, a aplicação composta de dois difeomorfismos é ainda um difeomorfismo.

Se $f\colon U \to V$ é um difeomorfismo, a Regra da Cadeia nos garante que sua derivada $f'(x)\colon \mathbb{R}^m \to \mathbb{R}^m$, em cada ponto $x \in U$, é um isomorfismo. Em termos do determinante jacobiano, isto significa que $\det Jf(x) \neq 0$ para todo $x \in U$, onde $Jf(x) = \left(\dfrac{\partial f_i}{\partial x_j}(x)\right)$ é a matriz jacobiana de f no ponto x. É natural indagar se vale a recíproca pois, afinal de contas, a utilidade das aplicações diferenciáveis está em que se podem deduzir algumas de suas propriedades a partir de propriedades das suas derivadas. Vejamos alguns exemplos.

Exemplo 11. Uma função diferenciável $f\colon I \to J$, de um intervalo $I \subset \mathbb{R}$ sobre o intervalo $J \subset \mathbb{R}$, é um difeomorfismo se, e somente se,

sua derivada $f'(x)$ é diferente de zero para todo $x \in I$. Com efeito, se $f'(x) \neq 0$ para todo $x \in I$ então, pelo Teorema de Darboux, ou $f'(x) > 0$ para todo $x \in I$, donde f é um homeomorfismo crescente, ou então $f'(x) < 0$ para todo $x \in I$, e f é um homeomorfismo decrescente. Em qualquer caso, pelo Teorema da Derivação da Função Inversa, (Volume 1, página 263) $f^{-1}\colon J \to I$ é diferenciável.

Exemplo 12. Seja $U \subset \mathbb{R}^m$ a bola aberta de centro na origem e raio 1. A aplicação $f\colon U \to \mathbb{R}^m$, definida por $f(x) = \dfrac{x}{\sqrt{1 - \langle x, x\rangle}}$, é um difeomorfismo de classe C^∞, cujo inverso $g\colon \mathbb{R}^m \to U$ é dado por $g(y) = \dfrac{y}{\sqrt{1 + \langle y, y\rangle}}$.

Exemplo 13. A aplicação $f\colon \mathbb{R}^2 \to \mathbb{R}^2$, dada por $f(x,y) = e^x(\cos y, \operatorname{sen} y)$ é de classe C^∞. Em termos de variável complexa, $f(z) = e^z$. Para todo $z = (x,y) \in \mathbb{R}^2$, a derivada $f'(z)\colon \mathbb{R}^2 \to \mathbb{R}^2$ é um isomorfismo, que consiste na multiplicação pelo número complexo não-nulo e^z. A matriz jacobiana de f no ponto $z = (x,y)$ tem a forma

$$Jf(x,y) = \begin{pmatrix} e^x \cos y & -e^x \operatorname{sen} y \\ e^x \operatorname{sen} y & e^x \cos y \end{pmatrix}$$

O determinante jacobiano de f é e^{2x}, donde $\neq 0$. Vemos, porém, que f não é injetiva: $f(x_1, y_1) = f(x_2, y_2) \Leftrightarrow x_1 = x_2$ e $y_2 = y_1 + 2k\pi$, $k \in \mathbb{Z}$. Geometricamente, f transforma cada reta vertical $x = a$ sobre o círculo de raio e^a e centro na origem, com período 2π. Cada reta horizontal $y = b$ é levada bijetivamente sobre a semi-reta aberta que parte da origem e passa pelo ponto $(\cos b, \operatorname{sen} b) \in \mathbb{R}^2$. Temos $f(\mathbb{R}^2) = \mathbb{R}^2 - \{0\}$.

Dizemos que uma aplicação diferenciável $f\colon U \to \mathbb{R}^m$, definida no aberto $U \subset \mathbb{R}^m$, é um *difeomorfismo local* quando para cada $x \in U$ existe um aberto V_x, com $x \in V_x \subset U$, tal que a restrição de f a V_x é um difeomorfismo sobre um aberto $W_x \subset \mathbb{R}^m$. Se $f \in C^k$ dizemos que f é um *difeomorfismo local de classe* C^k. Neste caso, para cada $x \in U$, a aplicação inversa $(f|V_x)^{-1}\colon W_x \to V_x$ também é de classe C^k. (Corolário 4 da Regra da Cadeia.)

Se $f\colon U \to \mathbb{R}^m$ é um difeomorfismo local de U sobre $V = f(U)$ então, para cada $x \in U$, a derivada $f'(x)\colon \mathbb{R}^m \to \mathbb{R}^m$ é um isomorfismo. O Teorema da Aplicação Inversa estabelecerá a recíproca, no caso em que $f \in C^k$ $(k \geq 1)$.

Pelo Exemplo 11 acima, todo difeomorfismo local $f\colon I \to \mathbb{R}$, definido num intervalo aberto $I \subset \mathbb{R}$, é um difeomorfismo (global) de I sobre sua imagem $J = f(I)$.

Todo difeomorfismo é, evidentemente, um difeomorfismo local.

Todo difeomorfismo local $f\colon U \to \mathbb{R}^m$ é uma aplicação aberta, isto é, transforma cada aberto $V \subset U$ num aberto $f(V) \subset \mathbb{R}^m$. Isto é claro se $V \subset V_x$, para algum $x \in U$, pois f é um homeomorfismo de V_x, sobre o aberto $W_x \subset \mathbb{R}^m$. No caso geral, temos $V = \bigcup_{x\in U} (V_x \cap V)$, donde $f(V) = \bigcup_{x\in U} f(V_x \cap V)$. Como cada $V_x \cap V$ é um aberto contido em V_x, sua imagem é aberta, logo $f(V)$ é uma reunião de conjuntos abertos, o que prova a afirmação feita.

Segue-se daí que um difeomorfismo local $f\colon U \to \mathbb{R}^m$ é um difeomorfismo (global) sobre sua imagem $V = f(U)$ se, e somente se, é uma aplicação injetiva.

A aplicação exponencial complexa, vista no Exemplo 13, é um difeomorfismo local, de classe C^∞, de $\mathbb{R}^2 - \{0\}$. Isto decorre imediatamente do Teorema da Aplicação Inversa, a ser demonstrado a seguir, ou da observação de que todo ramo de $\log w$ tal que $\log w_0 = z_0$ é um inverso local para f na vizinhança de z_0.

Basearemos a demonstração do Teorema da Aplicação Inversa no *método das aproximações sucessivas*, princípio de grande utilidade para provar a existência e a unicidade de soluções para equações diferenciais, equações integrais, etc.

Seja $X \subset \mathbb{R}^m$. Uma aplicação $f\colon X \to \mathbb{R}^m$ chama-se uma *contração* quando existem $\lambda \in \mathbb{R}$, $0 \leq \lambda < 1$, e uma norma em $\mathbb{R}^m$ relativamente à qual se tem $|f(x) - f(y)| \leq \lambda|x - y|$ para quaisquer $x, y \in X$. Toda contração é uniformemente contínua. Às vezes, quando for preciso especificar a constante λ, diremos que f é uma λ-*contração*.

Por exemplo, seja $U \subset \mathbb{R}^m$ aberto e convexo. Se $f\colon U \to \mathbb{R}^m$ é uma aplicação diferenciável tal que $|f'(x)| \leq \lambda < 1$ para uma certa constante λ e todo $x \in U$, a Desigualdade do Valor Médio dá $|f(x) - f(y)| \leq \lambda|x - y|$, logo f é uma contração.

Um *ponto fixo* de uma aplicação $f\colon X \to \mathbb{R}^m$ $(X \subset \mathbb{R}^m)$ é um ponto $x \in X$ tal que $f(x) = x$.

Teoremas que asseguram a existência de ponto fixo para certos tipos de aplicações são sempre interessantes pois, em princípio, a busca de uma solução x para uma equação do tipo $f(x) = b$ reduz-se à procura

de um ponto fixo para a aplicação ξ, definida por $\xi(x) = f(x) + x - b$. Com efeito, $\xi(x) = x \Leftrightarrow f(x) = b$.

Entre os vários teoremas de pontos fixos que se conhecem, o teorema abaixo, a despeito da simplicidade de sua demonstração, é um dos mais úteis, provando a existência de um ponto fixo, assegurando sua unicidade e fornecendo ainda um método iterativo para a obtenção de valores aproximados para o ponto fixo em questão.

Teorema do Ponto Fixo para Contrações (Método das aproximações sucessivas.) *Sejam $F \subset \mathbb{R}^m$ um subconjunto fechado e $f\colon F \to F$ uma contração. Dado qualquer ponto $x_0 \in F$, a seqüência $x_1 = f(x_0), x_2 = f(x_1), \ldots, x_{k+1} = f(x_k), \ldots$ converge para um ponto $a \in F$, que é o único ponto fixo de f.*

Demonstração: De $|f(x) - f(y)| \leq \lambda|x - y|$ segue-se que $|x_{k+1} - x_k| \leq \lambda^k|x_1 - x_0|$. Portanto

$$|x_{k+p} - x_k| \leq \sum_{i=0}^{p-1} |x_{k+i+1} - x_{k+i}| \leq \left[\sum_{i=0}^{p-1} \lambda^{k+i}\right] |x_1 - x_0| \leq$$
$$\leq \frac{\lambda^k}{1-\lambda} \cdot |x_1 - x_0|.$$

Como $0 \leq \lambda < 1$, temos $\lim\limits_{k\to\infty} |x_{k+p} - x_k| = 0$, seja qual for $p \in \mathbb{N}$. Logo (x_k) é uma seqüência de Cauchy em $\mathbb{R}^m$. Existe portanto $a = \lim x_k$ e $a \in F$ porque F é fechado. Como f é contínua, $f(a) = \lim\limits_{k\to\infty} x_{k+1} = a$, ou seja, a é um ponto fixo de f. Se tivermos ainda $f(b) = b$, com $b \in F$ então $|b - a| = |f(b) - f(a)| \leq \lambda|b - a|$, logo $(1 - \lambda)|b - a| \leq 0$. Como $1 - \lambda > 0$, isto dá $b = a$. Portanto, o ponto fixo de f é único.

Exemplo 14. Se tivermos apenas $|f(x) - f(y)| < |x - y|$ para todo par de pontos $x \neq y$ em F, o ponto fixo de $f\colon F \to F$ pode não existir. Consideremos, por exemplo, $f\colon \mathbb{R} \to \mathbb{R}$ dada por $f(x) = \frac{1}{2}\,(x+\sqrt{1+x^2})$. Tem-se $0 < f'(x) < 1$ para todo $x \in \mathbb{R}$, pois $f'(x) = \frac{1}{2}\left(1 + \frac{x}{\sqrt{1+x^2}}\right)\cdot$ Mas $f(x) > x$ para todo $x \in \mathbb{R}$, logo f não possui ponto fixo. Um exemplo mais simples (cujo domínio não é, entretanto, um subconjunto fechado) é o da função $g\colon (0,1) \to (0,1)$, definida por $g(x) = x^2/2$. Tem-se $0 < g'(x) < 1$ para todo $x \in (0,1)$ mas g não possui ponto fixo em $(0,1)$. Observe-se, finalmente, que se $X \subset \mathbb{R}^m$ é *compacto* e a

aplicação $f\colon X \to X$ cumpre a condição $|f(x) - f(y)| < |x - y|$ para todo par de pontos $x \neq y$ em X então f possui um ponto fixo único em X. Com efeito, seja $a \in X$ o ponto onde a função contínua $\varphi\colon X \to \mathbb{R}$, $\varphi(x) = |x - f(x)|$, atinge seu mínimo $c = |a - f(a)|$. Se fosse $c \neq 0$ então teríamos $a \neq f(a)$, donde $|f(a) - f(f(a))| < |a - f(a)|$ e daí o ponto $b = f(a)$ seria tal que $|b - f(b)| < c$, uma contradição. Logo $a = f(a)$. A unicidade é imediata: de $f(a) = a$ e $f(b) = b$ com $a \neq b$ resultaria $|a - b| = |f(a) - f(b)| < |a - b|$, um absurdo.

Para garantir que uma contração $f\colon X \to \mathbb{R}^m$ possui um ponto fixo $a \in X$, é essencial descobrir um subconjunto $F \subset X$ tal que $f(F) \subset F$ e F é fechado em $\mathbb{R}^m$. O lema abaixo é freqüentemente utilizado para esse fim.

Lema. *Seja $f\colon X \to \mathbb{R}^m$ uma λ-contração. Se X contém a bola fechada $B[a;r]$, tal que $|f(a) - a| \leq (1 - \lambda)r$, então f admite um ponto fixo em $B[a;r]$.*

Demonstração: Basta provar que $f(B[a;r]) \subset B[a;r]$. Ora,

$$x \in B[a;r] \Rightarrow |x - a| \leq r \Rightarrow |f(x) - a| \leq |f(x) - f(a)| + \\ + |f(a) - a| \leq \lambda|x - a| + (1 - \lambda)r \leq \lambda r + (1 - \lambda)r = r.$$

Portanto, $x \in B[a;r] \Rightarrow f(x) \in B[a;r]$, o que prova o lema.

O teorema seguinte e seu corolário são versões não-diferenciáveis do Teorema da Aplicação Inversa e, ao mesmo tempo, constituem etapas cruciais da sua demonstração. A parte mais importante do que eles afirmam é a que assegura ser a imagem $f(U)$ um conjunto aberto. Implícito nesta afirmação está um teorema de existência, em cuja prova utilizamos o método das aproximações sucessivas.

Teorema da Perturbação da Identidade. *Seja $\varphi\colon U \to \mathbb{R}^m$ uma contração definida no aberto $U \subset \mathbb{R}^m$. A aplicação $f\colon U \to \mathbb{R}^m$, dada por $f(x) = x + \varphi(x)$, é um homeomorfismo de U sobre o conjunto aberto $f(U) \subset \mathbb{R}^m$. Além disso, se $U = \mathbb{R}^m$ tem-se $f(U) = \mathbb{R}^m$.*

Demonstração: Para $x, y \in U$ quaisquer, temos

$$|f(x) - f(y)| = |x - y + \varphi(x) - \varphi(y)| \geq |x - y| - |\varphi(x) - \varphi(y)| \geq \\ \geq |x - y| - \lambda|x - y| = (1 - \lambda)|x - y|.$$

Daí resulta que f é uma bijeção de U sobre $f(U)$ e que a aplicação inversa $f^{-1}\colon f(U) \to U$ cumpre a condição de Lipschitz $|f^{-1}(w) - f^{-1}(z)| \leq$

$c|w-z|$, com $c = 1/(1-\lambda)$. Em particular, f é um homeomorfismo de U sobre $f(U)$. Para provar que $f(U) \subset \mathbb{R}^m$ é aberto, seja $b \in f(U)$. Temos $b = f(a) = a + \varphi(a)$ para um certo $a \in U$. Devemos mostrar que b é um ponto interior do conjunto $f(U)$, ou seja, que para todo ponto y suficientemente próximo de b, a equação $y = f(x)$ ou, o que é o mesmo, $y = x + \varphi(x)$, possui uma solução $x \in U$. Dentro do esquema de resolver equações por meio de pontos fixos, seja $r > 0$ tal que $B[a;r] \subset U$ e consideremos a aplicação $\xi = \xi_y \colon B[a;r] \to \mathbb{R}^m$, dada por $\xi(x) = y - \varphi(x)$. Então $\xi(x) = x \Leftrightarrow f(x) = y$. Como y é constante, ξ é uma contração. Pelo Lema acima, existe um ponto fixo $x \in B[a;r]$ para ξ desde que $|\xi(a) - a| \leq (1-\lambda)r$. Como $\xi(a) - a = y - \varphi(a) - a = y - b$, vemos que $\xi = \xi_y$ tem um ponto fixo em $B[a;r]$, (ou seja, existe $x \in B[a;r]$ tal que $f(x) = y$) contanto que tomemos $|y-b| \leq (1-\lambda)r$. Isto significa que, pondo $s = (1-\lambda)r$, temos $B(b;s) \subset f(B[a;r]) \subset f(U)$, logo $b \in \operatorname{int} f(U)$ e, como $b \in f(U)$ é arbitrário, o conjunto $f(U)$ é aberto. Finalmente, se $U = \mathbb{R}^m$ então, para todo $r > 0$ temos $B[a;r] \subset U$ donde $B[f(a);(1-\lambda)r] \subset f(U)$. Se tomarmos, para cada $k \in \mathbb{N}$, $r_k = k/(1-\lambda)$, veremos então que $B[f(a);k] \subset f(U)$, donde $\mathbb{R}^m = \bigcup\limits_{k\in\mathbb{N}} B[f(a);k] \subset f(U)$, ou seja, $f(U) = \mathbb{R}^m$.

Corolário (Perturbação de um isomorfismo.) *Sejam $U \subset \mathbb{R}^m$ aberto e $f \colon U \to \mathbb{R}^m$ uma aplicação da forma $f(x) = T\cdot x + \varphi(x)$, onde $T \colon \mathbb{R}^m \to \mathbb{R}^m$ é uma transformação linear invertível e $\varphi \colon U \to \mathbb{R}^m$ satisfaz $|\varphi(x) - \varphi(y)| \leq \lambda|x-y|$, com $\lambda \cdot |T^{-1}| < 1$. Então f é um homeomorfismo de U sobre o conjunto aberto $f(U) \subset \mathbb{R}^m$. Se $U = \mathbb{R}^m$, tem-se $f(U) = \mathbb{R}^m$.*

Com efeito, neste caso a aplicação $\psi = T^{-1} \circ \varphi \colon U \to \mathbb{R}^m$ cumpre

$$|\psi(x) - \psi(y)| = |T^{-1}[\varphi(x) - \varphi(y)]| \leq |T^{-1}|\,|\varphi(x) - \varphi(y)| \leq$$
$$\leq |T^{-1}| \cdot \lambda|x-y|.$$

Como $|T^{-1}| \cdot \lambda < 1$, vemos que $\psi = T^{-1} \circ \varphi$ é uma contração. Sendo $(T^{-1} \circ f)(x) = x + \psi(x)$, segue-se do teorema acima que $T^{-1} \circ f$ é um homeomorfismo de U sobre o aberto $T^{-1} \cdot f(U)$, donde f e um homeomorfismo de U sobre o aberto $f(U)$. Se $U = \mathbb{R}^m$ então, pelo teorema, $(T^{-1} \circ f)(U) = \mathbb{R}^m$, donde $f(U) = T \cdot \mathbb{R}^m = \mathbb{R}^m$.

Provaremos, a seguir, uma versão do Teorema da Aplicação Inversa na qual a aplicação é suposta diferenciável apenas num ponto. Na versão clássica, que decorre imediatamente desta, supõe-se que a aplicação seja pelo menos de classe C^1. Começamos com o

Lema (Diferenciabilidade do homeomorfismo inverso.) *Seja $f\colon U \to V$ um homeomorfismo entre os abertos $U, V \subset \mathbb{R}^m$. Se f é diferenciável num ponto $a \in U$ e a derivada $f'(a)\colon \mathbb{R}^m \to \mathbb{R}^m$ é um isomorfismo então o homeomorfismo inverso $f^{-1}\colon V \to U$ é diferenciável no ponto $b = f(a)$. Se f é fortemente diferenciável no ponto a então f^{-1} é fortemente diferenciável no ponto b.*

Demonstração: Escrevamos $g = f^{-1}$. Como o único candidato possível para derivada de g no ponto b é $f'(a)^{-1}$, escrevamos

$$(*) \qquad g(b+w) - g(b) = f'(a)^{-1} \cdot w + s(w)$$

e procuremos mostrar que $\lim_{w\to 0} s(w)/|w| = 0$. Ponhamos $v = g(b+w) - g(b)$. Então $f(a+v) - f(a) = f[g(b) + g(b+w) - g(b)] - b = f(g(b+w)) - b = b + w - b = w$. Como f e g são contínuas, $w \to 0$ se, e somente se, $v \to 0$. A diferenciabilidade de f no ponto a fornece

$$(**) \qquad f(a+v) - f(a) = f'(a) \cdot v + r(v), \text{ onde } \lim_{v\to 0} \frac{r(v)}{|v|} = 0.$$

Na igualdade (*), substituimos o primeiro membro por v e, no segundo, substituimos $w = f(a+v) - f(a)$ pelo segundo membro de (**). Resulta:

$$v = v + f'(a) \cdot r(v) + s(w),$$

donde

$$s(w) = -f'(a)^{-1} \cdot r(v) \text{ e } \frac{s(w)}{|w|} = -f'(a)^{-1} \cdot \frac{r(v)}{|v|} \cdot \frac{|v|}{|w|}\cdot$$

Quando $w \to 0$, tem-se também $v \to 0$, como vimos, logo $\lim_{w\to 0} \frac{r(v)}{|v|} = 0$. Além disso, pelo Teorema 3 (item 1º), o quociente $|v|/|w| = |v|/(|f(a+v) - f(a)|)$ é limitado nas proximidades de $v = 0$. Como a transformação linear $f'(a)^{-1}$ é contínua e se anula na origem, segue-se da expressão

$$\frac{s(w)}{|w|} = -f'(a)^{-1} \cdot \frac{r(v)}{|v|} \cdot \frac{|v|}{|w|}$$

que $\lim_{w\to 0} \frac{s(w)}{|w|} = 0$, donde $g = f^{-1}$ é diferenciável no ponto $b = f(a)$. Quanto à diferenciabilidade forte, se pusermos $v = g(b+w) - g(b)$ e $u = g(b+z) - g(b)$ resultará, como acima, que

$$s(w) - s(z) = f'(a)^{-1} \cdot [r(u) - r(v)].$$

Como toda transformação linear é Lipschitziana, a diferenciabilidade forte de g no ponto b resulta imediatamente da diferenciabilidade forte de f no ponto a, pelos Teoremas 4 e 5.

Teorema da Aplicação Inversa. *Sejam $f: U \to \mathbb{R}^m$, definida no aberto $U \subset \mathbb{R}^m$, fortemente diferenciável no ponto $a \in U$ e $f'(a): \mathbb{R}^m \to \mathbb{R}^m$ um isomorfismo. (Equivalentemente: o determinante jacobiano $\det Jf(a) = \det \left(\dfrac{\partial f_i}{\partial x_j}(a)\right)$ é diferente de zero.) Então f é um homeomorfismo de um aberto V contendo a sobre um aberto W contendo $f(a)$. O homeomorfismo inverso $f^{-1}: W \to V$ é fortemente diferenciável no ponto $f(a)$ e sua derivada nesse ponto é $[f'(a)]^{-1}$. Se $f \in C^k\,(k \geq 1)$ então V pode ser tomado de modo que f seja um difeomorfismo de V sobre W.* [Pelo Corolário 4 da Regra da Cadeia, tem-se ainda $f^{-1} \in C^k$.]

Demonstração: O trabalho está essencialmente feito. Resta apenas juntar as peças. Para simplificar a notação, suporemos $a = f(a) = 0$, o que não restringe a generalidade. Escrevendo $f(x) = f'(a) \cdot x + r(x)$, a diferenciabilidade forte de f assegura, via Teorema 5, que existe uma bola aberta V, de centro a, tal que $x, y \in V \Rightarrow |r(x) - r(y)| \leq \lambda \cdot |x - y|$, com $\lambda \cdot |f'(a)^{-1}| < 1$. Portanto, f é, em V, uma perturbação do isomorfismo $f'(a)$. Segue-se que f é um homeomorfismo de V sobre o aberto $W = f(V)$ e, pelo Lema acima, a inversa $f^{-1}: W \to V$ é fortemente diferenciável no ponto $f(a)$. No caso particular de $f \in C^k$ $(k \geq 1)$, a aplicação derivada $f': U \to \mathcal{L}(\mathbb{R}^m; \mathbb{R}^m)$ é contínua. Como o conjunto $\mathrm{GL}(\mathbb{R}^m)$ dos isomorfismos lineares de $\mathbb{R}^m$ é aberto em $\mathcal{L}(\mathbb{R}^m; \mathbb{R}^m)$ e $f'(a) \in \mathrm{GL}(\mathbb{R}^m)$, a bola V de centro a pode ser tomada, se necessário, tão pequena que $f'(x): \mathbb{R}^m \to \mathbb{R}^m$ seja ainda isomorfismo, para todo $x \in V$. Pelo lema da diferenciabilidade do homeomorfismo inverso, $f^{-1}: W \to V$ é diferenciável em todos os pontos de W, logo f é um difeomorfismo.

Corolário 1. *Uma aplicação $f: U \to \mathbb{R}^m$, de classe C^k no aberto $U \subset \mathbb{R}^m$, (onde $1 \leq k \leq \infty$) é um difeomorfismo local se, e somente se, para todo $x \in U$, a derivada $f'(x): \mathbb{R}^m \to \mathbb{R}^m$ é um isomorfismo (isto é, $\det Jf(x) \neq 0$).*

Corolário 2 (Perturbação diferenciável da identidade). *Seja $U \subset \mathbb{R}^m$ aberto e convexo. Se $\varphi: U \to \mathbb{R}^m$ é de classe C^1, com $|\varphi'(x)| \leq \lambda < 1$ para todo $x \in U$ e λ constante, então $f: U \to \mathbb{R}^m$, definida por $f(x) = x + \varphi(x)$, é um difeomorfismo de U sobre sua imagem $f(U)$. Se $U = \mathbb{R}^m$*

então $f(U) = \mathbb{R}^m$.

Com efeito, sendo U convexo, φ é uma contração, logo f é um homeomorfismo de U sobre o aberto $f(U)$, pelo Teorema da Perturbação da Identidade. Como $f'(x) = I + \varphi'(x)$ e $|\varphi'(x)| < 1$, segue-se que $f'(x)$ é, para todo $x \in U$, uma transformação linear invertível. Portanto $f^{-1} \colon f(U) \to U$ é diferenciável e f é um difeomorfismo.

Observação: A questão de saber em que condições um difeomorfismo local $f \colon U \to \mathbb{R}^m$ é, na realidade, um difeomorfismo global de U sobre $f(U)$, ou seja, quando é que f é injetivo, é um problema topológico. A teoria dos chamados "espaços de recobrimento"é muitas vezes útil nesse contexto, fornecendo respostas como, por exemplo, a seguinte: se U é um aberto limitado, $f(U)$ é simplesmente conexo e, para toda seqüência de pontos $x_k \to a \in \partial U$, a seqüência $(f(x_k))$ não possui subseqüência convergente para um ponto de $f(U)$, então o difeomorfismo local $f \colon U \to \mathbb{R}^m$ é injetivo, logo é um difeomorfismo de U sobre $f(U)$. Veja [10], pág. 122.

Exemplo 15. Seja $f \colon \mathbb{R}^{n^2} \to \mathbb{R}^{n^2}$ definida por $f(X) = X^k$ onde k é um número natural fixo e X^k é a k-ésima potência da matriz X, de n linhas e n colunas. A aplicação f é de classe C^∞ e sua derivada, em cada ponto X, é a transformação linear $f'(X) \colon \mathbb{R}^{n^2} \to \mathbb{R}^{n^2}$, dada por $f'(X) \cdot V = \sum_{i=1}^{k} X^{i-1} \cdot V \cdot X^{k-i}$. Por exemplo, se $k = 2$, temos $f'(X) \cdot V = V \cdot X + X \cdot V$. No ponto $X = I$ (matriz identidade), temos $f'(I) \cdot V = k \cdot V$, logo a derivada $f'(I) \colon \mathbb{R}^{n^2} \to \mathbb{R}^{n^2}$ é um isomorfismo. Pelo Teorema da Aplicação Inversa, existem abertos U, W em $\mathbb{R}^{n^2}$, ambos contendo $I = f(I)$, tais que $f \colon U \to W$ é um difeomorfismo C^∞. Isto significa que toda matriz Y suficientemente próxima da identidade ($Y \in W$) possui uma raiz k-ésima X (isto é, $X^k = Y$), a qual é única se impusermos que ela esteja suficientemente próxima da identidade ($Y \in U$). Além disso, X é uma função de classe C^∞ de Y.

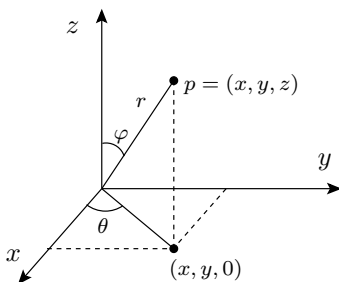
9 Aplicação: o Lema de Morse

Como ilustração sobre o emprego do Teorema da Aplicação Inversa, demonstraremos aqui o Lema de Morse; segundo ele, na vizinhança de um ponto crítico não-degenerado de uma função f, é possível tomar

um sistema de coordenadas em relação ao qual f se exprime como uma forma quadrática com coeficientes constantes: $f(y) = \Sigma\, a_{ij} y_i y_j$.

Um *sistema de coordenadas* (curvilíneas) de classe C^k num aberto $U \subset \mathbb{R}^n$ é um difeomorfismo $\xi \colon V \to U$, de classe C^k, definido num aberto $V \subset \mathbb{R}^n$. As coordenadas de um ponto $p \in U$ no sistema ξ são os números $y_1, \ldots, y_n$ tais que $y = (y_1, \ldots, y_n) \in V$ e $\xi(y) = p$.

Exemplo 16. Seja $\rho = \{(x, 0) \in \mathbb{R}^2; x \geq 0\}$. Em $U = \mathbb{R}^2 - \rho$ podemos introduzir um sistema de coordenadas $\xi \colon V \to U$, definido em $V = (0, +\infty) \times (0, 2\pi)$ por $\xi(r, \theta) = re^{i\theta} = (r \cos \theta, r \operatorname{sen} \theta)$. Se $p = (x, y) = \xi(r, \theta)$ então $r = \sqrt{x^2 + y^2}$ é a distância de p à origem e θ mede, em radianos, o ângulo de Op com o semi-eixo positivo das abcissas. Os números r, θ são chamados as *coordenadas polares* de $p = (x, y)$. Mais geralmente, se $\rho \subset \mathbb{R}^2$ é qualquer semi-reta fechada partindo da origem, tomando um vetor unitário $u = (\cos a, \operatorname{sen} a) \in \rho$ podemos definir um sistema de coordenadas polares $\xi \colon (0, +\infty) \times (a, a + 2\pi) \to U = \mathbb{R}^2 - \rho$ pela mesma fórmula $\xi(r, \theta) = re^{i\theta}$.

Exemplo 17. Outro exemplo bem conhecido é o sistema de *coordenadas esféricas* $\xi \colon V \to \mathbb{R}^3$, com $V = (0, +\infty) \times (0, 2\pi) \times (0, \pi)$ e $\xi(r, \theta, \varphi) = (r \operatorname{sen} \varphi \cos \theta, r \operatorname{sen} \varphi \operatorname{sen} \theta, r \cos \varphi)$. ξ é um difeomorfismo de V sobre $U = \mathbb{R}^3 - P$, onde $P = \{(x, 0, z) \in \mathbb{R}^3; x \geq 0\}$ é um semi-plano fechado vertical. Se $\xi(r, \theta, \varphi) = (x, y, z) = p$, os números r, θ, φ são chamados as coordenadas esféricas do ponto p; r é a distância de p à origem, φ mede, em radianos o ângulo do raio Op com o semi-eixo positivo dos z enquanto θ é a medida em radianos do ângulo que $(x, y, 0)$ faz com o semi-eixo positivo dos x.

A introdução de novos sistemas de coordenadas em regiões do espaço euclidiano tem por finalidade simplificar a descrição de certos conjuntos ou funções. Por exemplo, em coordenadas esféricas, a função $f(x, y, z) =$

$\sqrt{x^2+y^2+z^2}$ se torna $f(r,\theta,\varphi)=r$ e a esfera $x^2+y^2+z^2=c$ é descrita pela equação $r=\sqrt{c}$. O Lema de Morse diz que na vizinhança de um ponto crítico não-degenerado é possível obter um sistema de coordenadas que simplifica bastante a forma da função.

Lema de Morse. *Seja a um ponto crítico não-degenerado de uma função $f\colon U\to\mathbb{R}$ de classe C^k $(k\geq 3)$ num aberto $U\subset\mathbb{R}^n$. Existe um sistema de coordenadas $\xi\colon V\to W$, de classe C^{k-2}, com $a\in W\subset U$, $0\in V$ e $\xi(0)=a$, tal que*

$$f\,\xi(y)-f(a)=\sum_{i,j=1}^{m}a_{ij}y_iy_j$$

para todo $y=(y_1,\ldots,y_n)\in V$, onde

$$a_{ij}=\frac{1}{2}\cdot\frac{\partial^2 f}{\partial x_i\partial x_j}(a).$$

Demonstração: Para simplificar a notação, suporemos $a=0$ e $f(a)=0$. Seja W uma bola aberta de centro 0, contida em U. Pela fórmula de Taylor com resto integral,

$$x\in W\;\Rightarrow\;f(x)=\sum_{i,j=1}^{n}a_{ij}(x)x_ix_j\,,$$

com

$$a_{ij}(x)=\int_0^1(1-t)\,\frac{\partial^2 f}{\partial x_i\partial x_j}(tx)\,dt.$$

Como vemos, cada a_{ij} é uma função de classe C^{k-2}, definida na bola W. Para cada $x\in W$, a matriz $A(x)=(a_{ij}(x))$ é simétrica, em virtude do Teorema de Schwarz. Podemos escrever $f(x)=\langle A(x)\cdot x,x\rangle$ para todo $x\in W$. Seja $A_0=A(0)=\dfrac{1}{2}\cdot\left(\dfrac{\partial^2 f}{\partial x_i\partial x_j}(0)\right)$. Dizer que o ponto crítico 0 é não-degenerado significa afirmar que a matriz simétrica A_0 é invertível. Assim, para cada $x\in W$, podemos escrever $A(x)$ como produto de A_0 por uma matriz que depende de x em classe C^{k-2} e que para $x=0$ é a identidade. Pelo Exemplo 15 acima, podemos tomar o raio da bola W tão pequeno que $x\in W\Rightarrow A(x)=A_0\cdot B(x)^2$, onde $B\colon W\to\mathbb{R}^{n^2}$ é de classe C^{k-2}. Como A_0 e $A(x)$ (para todo $x\in W$) são simétricas tomando transpostas obtemos:

$$A=A_0\cdot B^2=(B^*)^2\cdot A_0\,,\text{ donde }B^2=A_0^{-1}(B^*)^2A_0=(A_0^{-1}\cdot B^*\cdot A_0)^2.$$

(Onde omitimos a variável x, por simplicidade.) Ainda pelo Exemplo 15, se o raio de W for suficientemente pequeno, $x \in W$ implicará que $B(x)$ e $A_0^{-1} \cdot B(x)^* \cdot A_0$ estão tão próximas da identidade que, por terem quadrados iguais, são elas iguais. Então $B = A_0^{-1} \cdot B^* \cdot A_0$, ou seja, $A_0 \cdot B = B^* \cdot A_0$ e daí $A = A_0 \cdot B^2 = B^* \cdot A_0 \cdot B$. Assim:

$$x \in W \Rightarrow f(x) = \langle A(x) \cdot x, x \rangle = \langle B(x)^* \cdot A_0 \cdot B(x) \cdot x, x \rangle = \\ = \langle A_0 \cdot B(x) \cdot x,\, B(x) \cdot x \rangle.$$

Mostraremos agora que, se o raio da bola W for tomado suficientemente pequeno, a aplicação $\varphi\colon W \to \mathbb{R}^n$, definida por $\varphi(x) = B(x) \cdot x$, é um difeomorfismo de classe C^{k-2} sobre sua imagem. Com efeito, para todo $x \in W$ e todo $v \in \mathbb{R}^n$, temos $\varphi'(x) \cdot v = \dfrac{\partial \varphi}{\partial v}(x) = \dfrac{\partial B}{\partial v}(x) \cdot x + B(x) \cdot v$. Para $x = 0$, vem: $\varphi'(0) \cdot v = B(0) \cdot v = v$, logo $\varphi'(0)\colon \mathbb{R}^n \to \mathbb{R}^n$ é a transformação linear identidade. Resulta então do Teorema da Aplicação Inversa que, se o raio de W for tomado suficientemente pequeno, obteremos um difeomorfismo $\varphi\colon W \to V$ de classe C^{k-2}, com $\varphi(x) = B(x) \cdot x$, $\varphi(0) = 0$ e $f(x) = \langle A_0 \cdot \varphi(x), \varphi(x) \rangle$ para todo $x \in W$. Então $\xi = \varphi^{-1}\colon V \to W$ é um sistema de coordenadas de classe C^{k-2} tal que, para todo $y \in V$, $\;f\xi(y) = \langle A_0 \cdot y, y \rangle = \sum\limits_{i,j=j}^{n} a_{ij} y_i y_j$.

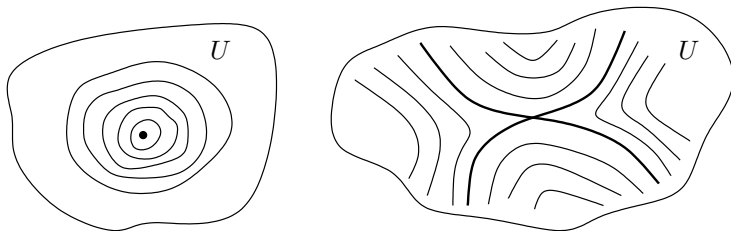
Corolário. *Nas condições do Lema de Morse, existe um sistema de coordenadas* $\zeta\colon V_0 \to W$, *de classe* C^{k-2}, *com* $a \in W \subset U$, $\;0 \in V_0$, $\zeta(0) = a$ *e* $f\zeta(y) - f(a) = -y_1^2 - \cdots - y_i^2 + y_{i+1}^2 + \cdots + y_n^2$.

É suficiente compor o sistema de coordenadas ξ com a mudança (linear) de coordenadas em $\mathbb{R}^n$ que torna a forma quadrática $\Sigma\, a_{ij} y_i y_j$ uma soma de quadrados, alguns dos quais precedidos do sinal menos. (Veja o Exemplo 18 do Capítulo III.) Este fato também decorre do Exemplo 29, Capítulo III, segundo o qual, dada a matriz simétrica $A_0 = (a_{ij})$, existe uma base ortonormal $\{u_1, \ldots, u_n\} \subset \mathbb{R}^n$ tal que $A_0 \cdot u_j = \lambda_j \cdot u_j \quad (j = 1, \ldots, n)$. Como A_0 é invertível, $\lambda_j \neq 0$ para todo j. Sejam $\lambda_1 < 0, \ldots, \lambda_i < 0$ e $\lambda_{i+1} > 0, \ldots, \lambda_n > 0$. Para $j \leq i$ ponhamos $v_j = u_j / \sqrt{-\lambda_j}$ e, para $j > i$ escrevamos $v_j = u_j / \sqrt{\lambda_j}$. Então $\{v_1, \ldots, v_n\}$ é ainda uma base de $\mathbb{R}^n$, valendo $\langle A_0 \cdot v_j, v_k \rangle = 0$ se $j \neq k$, $\;\langle A_0 \cdot v_j, v_j \rangle = -1$ se $j \leq i$ e $\langle A_0 \cdot v_j, v_j \rangle = 1$ se $j > i$. Seja agora $T\colon \mathbb{R}^n \to \mathbb{R}^n$ a transformação linear (invertível) tal que $T \cdot e_1 = v_1, \ldots, T \cdot e_n = v_n$. Pondo $V_0 = T^{-1}(V)$, onde V é o aberto

obtido no Lema de Morse, o difeomorfismo $\zeta = \xi \circ T \colon V_0 \to W$ cumpre:

$$f\zeta(y) = f\xi(T \cdot y) = \langle A_0 \cdot T \cdot y, T \cdot y\rangle = \langle A_0 \cdot \Sigma y_j v_j, \Sigma y_k v_k\rangle =$$
$$= \sum_{j,k} y_j y_k \langle A_0 \cdot v_j \, , v_k\rangle = -y_1^2 - \cdots - y_i^2 + y_{i+1}^2 + \cdots + y_n^2 \, .$$

O número i que aparece no Corolário acima chama-se o *índice* do ponto crítico a. Quando $i = n$, o ponto a é um máximo local para f; se $i = 0$, a é um ponto de mínimo local. Para $0 < i < n$, tem-se um *ponto de sela* de índice i.

Exemplo 18. Resulta do Corolário acima que, se $n = 2$, as curvas de nível na vizinhança de um ponto crítico não-degenerado de uma função $f \colon U \to \mathbb{R}$, definida num aberto do plano, tem uma das formas abaixo:

À esquerda, temos um ponto de máximo ou de mínimo; à direita um ponto de sela. Com efeito, se o ponto é de máximo ou de mínimo local, o Lema de Morse nos dá $f\zeta(y) - f(a) = \pm(y_1^2 + y_2^2)$, logo as curvas de nível de f próximas de a são imagens pelo difeomorfismo ζ dos círculos $y_1^2 + y_2^2 =$ constante, logo têm a forma dada pela figura à esquerda. A figura à direita, por sua vez, é a imagem por um difeomorfismo das curvas de nível da função $f\zeta(y) = -y_1^2 + y_2^2$.

Os três parágrafos seguintes têm objetivo semelhante ao deste: a partir de hipóteses de "regularidade"sobre a derivada, obter sistemas de coordenadas convenientes, em relação aos quais a aplicação se exprime por meio de fórmulas simples. Um problema análogo ocupa uma posição central em Álgebra Linear: dada uma transformação linear $T \colon \mathbb{R}^n \to \mathbb{R}^n$, achar uma base de $\mathbb{R}^n$ segundo a qual a matriz de T seja a mais simples. Quando é possível, tenta-se diagonalizar a matriz de T, isto é, obter uma base formada por vetores próprios. (No Capítulo III, §10, mostramos

que isto pode ser feito quando T é auto-adjunta.) Em geral, o melhor que se pode obter é uma base que dê a T uma "matriz de Jordan". No nosso caso, o Lema de Morse, as Formas Locais de Imersões, Submersões e o Teorema do Posto (estes três a serem provados adiante) constituem os casos mais elementares de formas locais de aplicações diferenciáveis. O prosseguimento desses estudos é feito na teoria das singularidades de aplicações diferenciáveis, iniciada por Morse, Whitney e Thom, e desenvolvida por vários autores em anos recentes.

10 A forma local das imersões

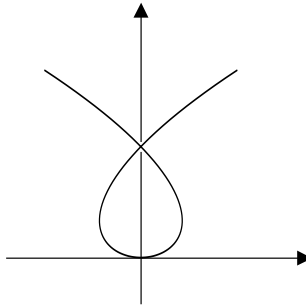
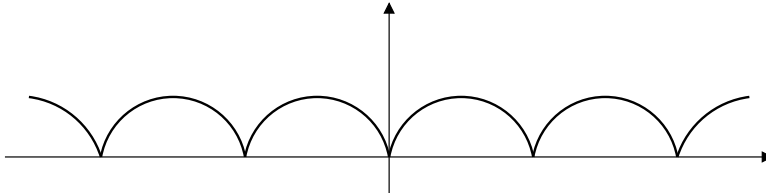
Uma *imersão* do aberto $U \subset \mathbb{R}^m$ no espaço euclidiano $\mathbb{R}^n$ é uma aplicação diferenciável $f\colon U \to \mathbb{R}^n$ tal que, para cada $x \in U$, a derivada $f'(x)\colon \mathbb{R}^m \to \mathbb{R}^n$ é uma transformação linear injetiva. Evidentemente, isto só pode ocorrer quando $m \leq n$.

A aplicação composta de duas imersões é ainda uma imersão.

Se a derivada $f'\colon U \to \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$ for contínua no ponto $a \in U$, fato que equivale a dizer que f é fortemente diferenciável nesse ponto, então a condição de $f'(a)\colon \mathbb{R}^m \to \mathbb{R}^n$ ser injetiva implica que f é injetiva numa vizinhança de a (Teorema 4) e, mais precisamente, f transforma uma bola de centro a homeomorficamente sobre sua imagem.

Exemplo 19. Seja $f\colon \mathbb{R}^m \to \mathbb{R}^m \times \mathbb{R}^n$ a aplicação de inclusão, dada por $f(x) = (x, 0)$. Como f é linear, temos $f'(x) = f$ para todo $x \in \mathbb{R}^m$, logo f é uma imersão C^∞. O principal resultado deste parágrafo mostra que toda imersão de classe C^k $(k \geq 1)$ se comporta localmente como esta.

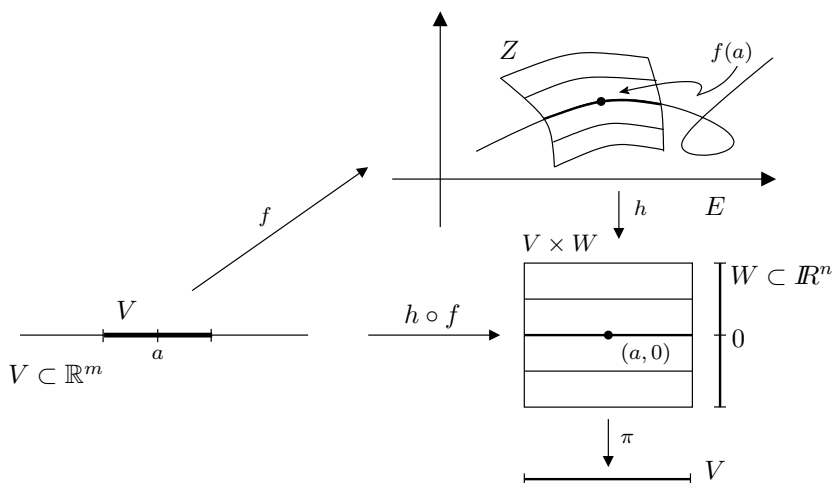
Exemplo 20. Seja $J \subset \mathbb{R}$ um intervalo aberto. Um caminho diferenciável $f\colon J \to \mathbb{R}^n$ é uma imersão se, e somente se, seu vetor-velocidade é diferente de zero em cada ponto $t \in J$. Isto significa que a imagem $f(J)$ possui, em cada ponto $f(t)$, uma reta tangente, a saber, a reta $L = \{f(t) + s \cdot f'(t); s \in \mathbb{R}\}$. Notemos que uma imersão pode não ser injetiva. Então, quando $f(t_1) = f(t_2)$, as duas retas tangentes $L_1 = f(t_1) + \mathbb{R} \cdot f'(t_1)$ e $L_2 = f(t_2) + \mathbb{R} \cdot f'(t_2)$ podem (ou não) ser distintas. Se considerarmos, entretanto, uma pequena vizinhança V de t_1 , nenhum outro ponto de V terá a mesma imagem que t_1 (Teorema 3) e assim L_1 será a única reta tangente no ponto $f(t_1)$ para o caminho restrito $f|V$. Um exemplo simples e elucidativo de imersão da reta no plano é dado

por $f\colon \mathbb{R} \to \mathbb{R}^2$, $f(t) = (t^3 - t, t^2)$. A imagem de f é esboçada na figura abaixo.

Exemplo 21. O caminho $g\colon \mathbb{R} \to \mathbb{R}^2$, definido por $g(t) = (t - \operatorname{sen} t, 1 - \cos t)$, é injetivo, de classe C^∞, mas não é uma imersão de $\mathbb{R}$ em $\mathbb{R}^2$. Sua imagem, uma curva chamada *ciclóide*, exibe uma infinidade de pontos angulares (cúspides), nos quais o vetor-velocidade $g'(t)$ deve forçosamente ser igual a zero.

Deve-se notar, entretanto, que os pontos onde a derivada de uma aplicação deixa de ser injetiva nem sempre se denunciam pela forma geométrica da imagem. Por exemplo, a imagem do caminho $f\colon \mathbb{R} \to \mathbb{R}^2$, $f(t) = (t^3, t^3)$, é uma reta. Para $t = 0$, o vetor-velocidade $f'(t)$ se anula, o que não se deve ao aspecto de $f(\mathbb{R})$ mas à maneira como tal conjunto está parametrizado por f.

O teorema seguinte mostra que, se a imersão $f\colon U \to \mathbb{R}^n$ é suficientemente suave, é possível introduzir novas coordenadas na vizinhança de cada ponto da imagem de modo que f assuma localmente a forma do Exemplo 19 acima.

Forma Local das Imersões. *Seja $f\colon U \to \mathbb{R}^{m+n}$ definida no aberto $U \subset \mathbb{R}^m$, fortemente diferenciável no ponto $a \in U$. Se a derivada $f'(a)\colon \mathbb{R}^m \to \mathbb{R}^{m+n}$ é injetiva, existe um homeomorfismo $h\colon Z \to V \times W$, de um aberto $Z \ni f(a)$ em $\mathbb{R}^{m+n}$ sobre um aberto $V \times W \ni (a, 0)$ em*

$\mathbb{R}^m \times \mathbb{R}^n$, *tal que* $hf(x) = (x, 0)$ *para todo* $x \in V$ *e* h *é fortemente diferenciável no ponto* $f(a)$. *Caso* $f \in C^k$ $(k \geq 1)$, *é possível restringir* V, W *e* Z *se necessário, de modo que* h *seja um difeomorfismo de classe* C^k.

A demonstração faz uso, de modo simples, do Teorema da Aplicação Inversa. Pelo menos tão importante quanto ela é a figura abaixo, que deve ser bem entendida pois contém o significado do teorema.

Demonstração: Seja $E = f'(a) \cdot \mathbb{R}^m$ a imagem de $f'(a)$. Existem vetores linearmente independentes $v_1, \ldots, v_n \in \mathbb{R}^{m+n}$, que geram um subespaço vetorial $F \subset \mathbb{R}^{m+n}$ tal que $\mathbb{R}^{m+n} = E \oplus F$. Definamos $\varphi \colon U \times \mathbb{R}^n \to \mathbb{R}^{m+n}$ pondo

$$\varphi(x, y) = f(x) + \sum_{i=1}^{n} y_i v_i \,,$$

onde $y = (y_1, \ldots, y_n)$. Então φ é fortemente diferenciável no ponto $(a, 0)$, e

$$\varphi'(a, 0) \cdot (v, w) = f'(a) \cdot v + \sum_{i=1}^{n} \beta_i v_i \,,$$

onde $v \in \mathbb{R}^m$ e $w = (\beta_1, \ldots, \beta_n) \in \mathbb{R}^n$. Sendo $f'(a)$ injetiva, e $\mathbb{R}^{m+n}$ a soma direta da imagem de $f'(a)$ com o subespaço gerado por $v_1, \ldots, v_n$, resulta imediatamente que $\varphi'(a, 0) \colon \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^{m+n}$ é injetiva e portanto um isomorfismo. Pelo Teorema da Aplicação Inversa, φ aplica

homeomorficamente um aberto contendo $(a, 0)$ (o qual podemos supor da forma $V \times W$, onde $W \ni 0$ em $\mathbb{R}^n$) sobre um aberto $Z \ni f(a)$ em $\mathbb{R}^{m+n}$. Seja $h = \varphi^{-1}: Z \to V \times W$. Como $\varphi(x, 0) = f(x)$ temos $hf(x) = h\varphi(x, 0) = (x, 0)$. Quando $f \in C^k$ $(k \geq 1)$ então φ também é de classe C^k. Ainda pelo Teorema da Aplicação Inversa, V e W podem ser tomados de modo que $\varphi|(V \times W)$ seja um difeomorfismo sobre $Z = \varphi(v \times W)$, cujo inverso h é também de classe C^k.

Corolário. *Seja $f: U \to \mathbb{R}^{m+n}$ definida no aberto $U \subset \mathbb{R}^m$, fortemente diferenciável no ponto a, com $f'(a): \mathbb{R}^m \to \mathbb{R}^{m+n}$ injetiva. Existe um aberto V, com $a \in V \subset U$, que é aplicado homeomorficamente por f sobre $f(V)$. O homeomorfismo inverso $f^{-1}: f(V) \to V$ é a restrição de uma aplicação contínua $\xi: Z \to V$, definida num aberto $Z \ni f(a)$ em $\mathbb{R}^{m+n}$ e ξ é fortemente diferenciável no ponto $f(a)$. Se $f \in C^k$ $(k \geq 1)$ então ξ pode ser tomada de classe C^k.*

Basta tomar $h: Z \to V \times W$ como no teorema acima e definir $\xi = \pi \circ h$, onde $\pi: V \times W \to V$ é a projeção sobre a primeira coordenada. Para todo $x \in V$, tem-se então $\xi(f(x)) = \pi h f(x) = \pi(x, 0) = x$, logo $\xi|f(V) = f^1$.

Observação: Decorre deste corolário que se $f \in C^k$ $(k \geq 1)$ e $f'(a)$ é injetiva para algum $a \in U$, então $f'(x): \mathbb{R}^m \to \mathbb{R}^{m+n}$ é injetiva para todo x num aberto V contendo a em $\mathbb{R}^m$. Com efeito, de $\xi f(x) = x$ para todo $x \in V$, resulta pela Regra da Cadeia que $\xi'(f(x)) \circ f'(x) = \text{Id}: \mathbb{R}^m \to \mathbb{R}^m$, logo $f'(x)$ é injetiva. Isto também pode ser verificado diretamente, como conseqüência dos seguintes fatos: 1º) A aplicação derivada $f': U \to \mathcal{L}(\mathbb{R}^m; \mathbb{R}^{m+n})$ é contínua; 2º) O conjunto das transformações lineares injetivas é aberto em $\mathcal{L}(\mathbb{R}^m; \mathbb{R}^{m+n}$. [Para comprovar o segundo fato, note que uma transformação linear T é injetiva se, e somente se, sua matriz contém um subdeterminante menor $m \times m$ não-nulo. O mesmo menor, sendo uma função contínua, será não-nulo numa pequena bola de centro T, logo todas as transformações pertencentes a essa bola serão injetivas.]

11 A forma local das submersões

Uma aplicação diferenciável $f: U \to \mathbb{R}^n$, definida num aberto $U \subset \mathbb{R}^m$, chama-se uma *submersão* quando, para todo $x \in U$, sua derivada

$f'(x)\colon \mathbb{R}^m \to \mathbb{R}^n$ é uma transformação linear sobrejetiva. Para que isto ocorra, é necessário que se tenha $m \geq n$.

Um funcional linear é sobrejetivo ou é nulo. Portanto uma função diferenciável $f\colon U \to \mathbb{R}$ é uma submersão se, e somente se, $df(x) \neq 0$ ou, equivalentemente, $\operatorname{grad} f(x) \neq 0$, para todo $x \in U$.

Segue-se da Regra da Cadeia que a aplicação composta de duas submersões é ainda uma submersão.

Uma "decomposição em soma direta do tipo $\mathbb{R}^{m+n} = \mathbb{R}^m \oplus \mathbb{R}^n$ "significa que se fez uma partição $\{1, \ldots, m+n\} = I \cup J$, onde $I = \{i_1, \ldots, i_m\}$ e $J = \{j_1, \ldots, j_n\}$ são disjuntos. Dada essa partição, consideramos $\mathbb{R}^m \subset \mathbb{R}^{m+n}$ como o subespaço gerado pelos vetores $\{e_{i_1}, \ldots, e_{i_m}\}$ e $\mathbb{R}^n \subset \mathbb{R}^{m+n}$ como o subespaço gerado por $\{e_{j_1}, \ldots, e_{j_n}\}$. Evidentemente, todo vetor $z \in \mathbb{R}^{m+n}$ se escreve, de modo único, como $z = x + y$ onde $x \in \mathbb{R}^m$ e $y \in \mathbb{R}^n$. (Substitua por zero cada coordenada de z cujo índice pertence a J e obtenha x; faça o mesmo com I em vez de J e obtenha y.) Assim, $\mathbb{R}^{m+n}$ é soma direta desses dois subespaços e escrevemos $\mathbb{R}^{m+n} = \mathbb{R}^m \oplus \mathbb{R}^n$. Seria bem mais preciso escrever $\mathbb{R}^m_I$ e $\mathbb{R}^n_J$ para indicar os subespaços acima, mas não há perigo em usar a notação mais simples, desde que mantenhamos em mente a partição dos índices. Uma vez dada a decomposição em soma direta $\mathbb{R}^{m+n} = \mathbb{R}^m \oplus \mathbb{R}^n$, escreveremos os elementos de $\mathbb{R}^{m+n}$ como pares $z = (x, y)$, $x \in \mathbb{R}^m$, $y \in \mathbb{R}^n$. Por exemplo, seja $\mathbb{R}^3 = \mathbb{R}^2 \oplus \mathbb{R}$, onde $\mathbb{R}^2$ é gerado por $\{e_1, e_3\}$ e $\mathbb{R}$ por $\{e_2\}$. Então todo $z = (x_1, x_2, x_3) \in \mathbb{R}^3$ será escrito como $z = (x, y)$, onde $x = (x_1, 0, x_3)$ e $y = (0, x_2, 0)$.

Dada uma transformação linear sobrejetiva $T\colon \mathbb{R}^{m+n} \to \mathbb{R}^n$, pode-se obter uma decomposição em soma direta do tipo $\mathbb{R}^{m+n} = \mathbb{R}^m \oplus \mathbb{R}^n$ tal que a restrição $T|\mathbb{R}^n\colon \mathbb{R}^n \to \mathbb{R}^n$ seja um isomorfismo. Basta observar que os vetores $T \cdot e_1, \ldots, T \cdot e_{m+n}$ geram $\mathbb{R}^n$, portanto é possível selecionar dentre eles uma base $\{T \cdot e_{j_1}, \ldots, T \cdot e_{j_n}\}$. Seja $J = \{j_1, \ldots, j_n\}$ e seja $I = \{i_1, \ldots, i_m\}$ o conjunto dos índices restantes. A partição $\{1, \ldots, m+n\} = I \cup J$ fornece a decomposição em soma direta $\mathbb{R}^{m+n} = \mathbb{R}^m \oplus \mathbb{R}^n$. A restrição $T|\mathbb{R}^n$ é um isomorfismo porque transforma uma base numa base. Seja (a_{ij}) a matriz $n \times (m+n)$ da transformação T em relação às bases canônicas. Dizer que $T|\mathbb{R}^n$ é um isomorfismo equivale a afirmar que a submatriz $n \times n$, obtida da anterior selecionando-se as n colunas cujos índices pertencem ao conjunto J, tem determinante $\neq 0$.

Exemplo 22. Dada uma decomposição em soma direta do tipo $\mathbb{R}^{m+n} = \mathbb{R}^m \oplus \mathbb{R}^n$, seja $f\colon \mathbb{R}^{m+n} \to \mathbb{R}^n$ a projeção sobre a segunda coordenada,

isto é, $f(x,y) = y$. Como f é linear, temos $f'(z) = f$ para todo $z \in \mathbb{R}^{m+n}$, logo f é uma submersão. A matriz jacobiana de f tem como linhas os n últimos vetores da base canônica de $\mathbb{R}^{m+n}$.

Seja $f\colon U \to \mathbb{R}^n$ definida no aberto $U \subset \mathbb{R}^m$. Seja ainda $E \subset \mathbb{R}^m$ um subespaço vetorial. Para todo ponto $a \in U$, podemos indagar se f é *diferenciável ao longo de* E no ponto a. Responder afirmativamente a esta pergunta significa dizer que existe uma transformação linear $\partial_E f(a)\colon E \to \mathbb{R}^n$, chamada a *derivada de* f *ao longo de* E no ponto a, tal que

$$v \in E, a + v \in U \Rightarrow f(a+v) - f(a) = \partial_E f(a) \cdot v + r(v),$$

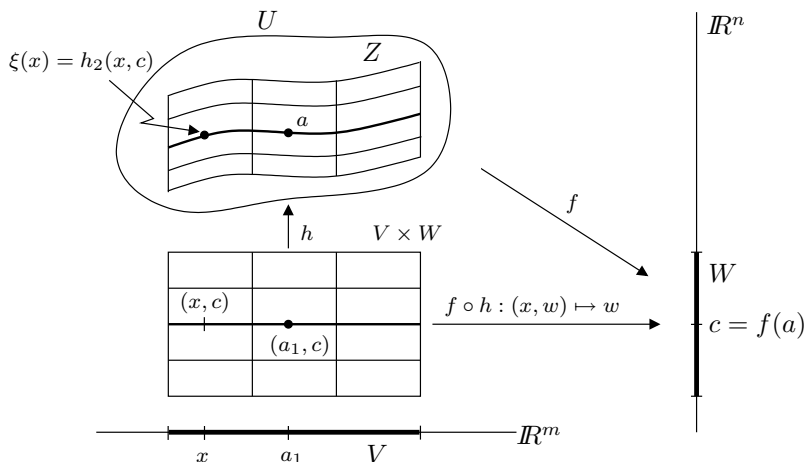
$$\text{com } \lim_{\substack{v\in E\\ v\to 0}} \frac{r(v)}{|v|} = 0.$$

Evidentemente, se f é diferenciável no ponto a, então f é diferenciável nesse ponto ao longo de qualquer subespaço $E \subset \mathbb{R}^m$, com $\partial_E f(a) = f'(a)|E =$ restrição de $f'(a)$ ao subespaço E.

Se é dada uma decomposição em soma direta do tipo $\mathbb{R}^{m+n} = \mathbb{R}^m \oplus \mathbb{R}^n$ e $f\colon U \to \mathbb{R}^p$ é definida num aberto $U \subset \mathbb{R}^{m+n}$, a derivada de f no ponto a ao longo de $\mathbb{R}^m$ é indicada por $\partial_1 f(a)$ e a derivada ao longo de $\mathbb{R}^n$ é representada por $\partial_2 f(a)$. Estas são as *derivadas parciais* de f no ponto a relativamente à decomposição $\mathbb{R}^{m+n} = \mathbb{R}^m \oplus \mathbb{R}^n$. Evidentemente, qualquer dessas derivadas parciais pode existir ou não. Quando existem, são transformações lineares $\partial_1 f(a)\colon \mathbb{R}^m \to \mathbb{R}^p$ e $\partial_2 f(a)\colon \mathbb{R}^n \to \mathbb{R}^p$. Se $f\colon U \to \mathbb{R}^p$ é diferenciável no ponto a então $\partial_1 f(a) = f'(a)|\mathbb{R}^m$ e $\partial_2 f(a) = f'(a)|\mathbb{R}^n$ são meramente as restrições da derivada $f'(a)$ aos subespaços que determinam a decomposição. Em particular, para qualquer $u = (v,w) \in \mathbb{R}^m \oplus \mathbb{R}^n$, $f'(a) \cdot u = f'(a) \cdot v + f'(a) \cdot w = \partial_1 f(a) \cdot v + \partial_1 f(a) \cdot w$. Sabemos porém que, mesmo no caso da decomposição usual $\mathbb{R}^2 = \mathbb{R} \oplus \mathbb{R}$, uma função $f\colon \mathbb{R}^2 \to \mathbb{R}$ pode ser diferenciável ao longo de cada um dos subespaços $\mathbb{R}$ sem ser diferenciável em $\mathbb{R}^2$.

O teorema abaixo mostra que o Exemplo 22 é, localmente, o caso mais geral de uma submersão. Ele diz que, dada uma submersão f de classe C^1, é possível tomar novas coordenadas em torno de cada ponto do seu domínio de modo que f seja a projeção sobre as n últimas coordenadas. Como no caso da forma local das imersões, é importante identificar o teorema com a figura que o acompanha.

Forma Local das Submersões. *Seja* $f\colon U \to \mathbb{R}^n$ *definida no aberto* $U \subset \mathbb{R}^{m+n}$ *e fortemente diferenciável no ponto* $a \in U$. *Se* $f'(a)\colon \mathbb{R}^{m+n} \to$

$\mathbb{R}^n$ é sobrejetiva ou, mais precisamente, se é dada uma decomposição em soma direta do tipo $\mathbb{R}^{m+n} = \mathbb{R}^m \oplus \mathbb{R}^n$ tal que $a = (a_1, a_2)$ e a derivada parcial $\partial_2 f(a) = f'(a)|\mathbb{R}^n \colon \mathbb{R}^n \to \mathbb{R}^n$ é um isomorfismo, então existem abertos, V, W, Z, com $a \in Z \subset \mathbb{R}^{m+n}$, $a_1 \in V \subset \mathbb{R}^m$, $f(a) \in W \subset \mathbb{R}^n$, e um homeomorfismo $h \colon V \times W \to Z$, fortemente diferenciável no ponto $(a_1, f(a))$, tal que $fh(x, w) = w$ para todo $(x, w) \in V \times W$. Se f é de classe C^k em U $(k \geq 1)$, podemos restringir V, W, Z, se necessário, de modo que h seja um difeomorfismo de classe C^k.

Demonstração: Seja $c = f(a)$. A aplicação $\varphi \colon U \to \mathbb{R}^m \times \mathbb{R}^n$, definida por $\varphi(x, y) = (x, f(x, y))$, é fortemente diferenciável no ponto a, e sua derivada $\varphi'(a) \colon \mathbb{R}^{m+n} \to \mathbb{R}^m \times \mathbb{R}^n$ é dada por $\varphi'(a) \cdot (v, w) = (v, f'(a) \cdot (v, w)) = (v, \partial_1 f(a) \cdot v + \partial_2 f(a) \cdot w)$. Ora, a transformação linear $(\alpha, \beta) \mapsto (\alpha, [\partial_2 f(a)]^{-1} \cdot (\beta - \partial_1 f(a) \cdot \alpha))$ é evidentemente uma inversa para $\varphi'(a)$. Logo $\varphi'(a)$ é um isomorfismo. Pelo Teorema da Aplicação Inversa, φ é um homeomorfismo (com inverso fortemente diferenciável no ponto $(a_1, c) = \varphi(a)$) de um aberto contendo a sobre um aberto contendo (a_1, c), o qual pode ser tomado da forma $V \times W$, com V aberto em $\mathbb{R}^m$ e W aberto em $\mathbb{R}^n$. Ponhamos $Z = \varphi^{-1}(V \times W)$ e $h = \varphi^{-1} \colon V \times W \to Z$. Como o homeomorfismo φ deixa fixa a primeira coordenada, seu inverso h tem a mesma propriedade: $h(x, w) = (x, h_2(x, w))$ para todo $(x, w) \in V \times W$. Então

$$(x, w) \in V \times W \Rightarrow (x, w) = \varphi h(x, w) = \varphi(x, h_2(x, w)) = \\ = (x, f(x, h_2(x, w))) = (x, fh(x, w)),$$

donde $fh(x, w) = w$ para todo $(x, w) \in V \times W$. Quando $f \in C^k$ então $\varphi \in C^k$. Ainda pelo Teorema da Aplicação Inversa, V, W e Z podem ser tomadas de modo que φ seja um difeomorfismo de classe C^k de Z sobre $V \times W$ e seu inverso h também é um difeomorfismo de classe C^k.

Corolário 1. *Seja $f\colon U \to \mathbb{R}^n$, definida no aberto $U \subset \mathbb{R}^{m+n}$, fortemente diferenciável no ponto $a \in U$. Se $f'(a)\colon \mathbb{R}^{m+n} \to \mathbb{R}^n$ é sobrejetiva, existe um aberto Z contendo a em $\mathbb{R}^{m+n}$ tal que $f|Z$ é uma aplicação aberta. (Ou seja, para todo aberto $A \subset Z$, $f(A)$ é aberto em $\mathbb{R}^n$.)*

Corolário 2. *Toda submersão de classe C^k, $(k \geq 1)$ é uma aplicação aberta.*

Os corolários resultam do teorema em virtude de ser a projeção $\pi\colon V \times W \to W$, na segunda coordenada, uma aplicação aberta. Com efeito, temos $f \circ h = \pi\colon V \times W \to W$. Logo, para todo $A \subset Z$ aberto, $f(A) = fhh^{-1}(A) = \pi h^{-1}(A)$. Como h é contínua, $h^{-1}(A)$ é aberto, logo sua projeção $\pi h^{-1}(A)$, ou seja, $f(A)$, é um conjunto aberto.

Observação: Na decomposição $\mathbb{R}^{m+n} = \mathbb{R}^m \oplus \mathbb{R}^n$, a que se refere o teorema acima, seja $\mathbb{R}^n$ gerado pelos vetores e_j, $j \in J = \{j_1, \ldots, j_n\}$, da base canônica de $\mathbb{R}^{m+n}$. A derivada parcial $\partial_2 f(a)\colon \mathbb{R}^n \to \mathbb{R}^n$ é um isomorfismo se, e somente se, a matriz $\left(\dfrac{\partial f_i}{\partial x_j}(a)\right)$, $1 \leq i \leq n$, $j \in J$, obtida da matriz jacobiana de f escolhendo as n colunas cujos índices pertencem ao conjunto J, tem determinante $\neq 0$.

Teorema da Aplicação Implícita. *Seja $f\colon U \to \mathbb{R}^n$, definida no aberto $U \subset \mathbb{R}^{m+n}$, fortemente diferenciável no ponto $a \in U$, com $f(a) = c$. Se $f'(a)\colon \mathbb{R}^{m+n} \to \mathbb{R}^n$ é sobrejetiva ou, mais precisamente, se $\mathbb{R}^{m+n} = \mathbb{R}^m \oplus \mathbb{R}^n$ é uma decomposição em soma direta tal que $a = (a_1, a_2)$ e a derivada $\partial_2 f(a)\colon \mathbb{R}^n \to \mathbb{R}^n$ é um isomorfismo, então existem abertos V, Z (onde $a_1 \in V \subset \mathbb{R}^m$, $a \in Z \subset U$) com a seguinte propriedade: para cada $x \in V$ há um único $\xi(x) \in \mathbb{R}^n$ tal que $(x, \xi(x)) \in Z$ e $f(x, \xi(x)) = c$. A aplicação $\xi\colon V \to \mathbb{R}^n$ assim definida é fortemente diferenciável no ponto a_1 e sua derivada nesse ponto é $\xi'(a_1) = -[\partial_2 f(a)]^{-1} \circ [\partial_1 f(a)]$. Se $f \in C^k$ $(k \geq 1)$, então $\xi \in C^k$ e sua derivada num ponto $x \in V$ qualquer é*

$$\xi'(x) = -[\partial_2 f(x, \xi(x))]^{-1} \circ [\partial_1 f(x, \xi(x))].$$

Em resumo (e mais precisamente): *$f^{-1}(c) \cap Z$ é o gráfico da aplicação $\xi\colon V \to \mathbb{R}^n$, fortemente diferenciável no ponto a_1. Se $f \in C^k$ então $\xi \in C^k$.*

A aplicação ξ diz-se *definida implicitamente* pela equação $f(x,y)=c$.

Demonstração: Usando a notação do teorema anterior, temos $Z = h(V\times W)$ e $h(x,w)=(x,h_2(x,w))$ para $(x,w)\in V\times W$. Defina $\xi\colon V\to\mathbb{R}^n$ pondo $\xi(x)=h_2(x,c)$. Então $(x,\xi(x))=h(x,c)\in Z$ para todo $x\in V$ e $f(x,\xi(x))=fh(x,c)=c$. Reciprocamente, se $(x,y)\in Z$ e $f(x,y)=c$ então $(x,y)=h\varphi(x,y)=h(x,c)=(x,h_2(x,c))=(x,\xi(x))$, logo $y=\xi(x)$. Isto mostra que, para cada $x\in V$, existe um único $\xi(x)\in\mathbb{R}^n$ tal que $(x,\xi(x))\in Z$ e $f(x,\xi(x))=c$. A aplicação $\xi\colon V\to\mathbb{R}^n$, assim definida, é dada por $\xi(x)=h_2(x,c)$, logo é fortemente diferenciável se f o é, e pertence à classe C^k se $f\in C^k$. Finalmente, derivando a igualdade $f(x,\xi(x))=c$ obtemos, pela Regra da Cadeia: $\partial_1 f(x,\xi(x))+\partial_2 f(x,\xi(x))\circ\xi'(x)=0$, donde $\xi'(x)=-[\partial_2 f(x,\xi(x))]^{-1}\circ[\partial_1 f(x,\xi(x))]$, quando $f\in C^k$ $(k\geq 1)$. Se f é suposta diferenciável apenas no ponto a, podemos ainda derivar a identidade $f(x,\xi(x))=c$ no ponto $x=a_1$, usando a Regra da Cadeia. Observando que $(a_1,\xi(a_1))=h(a_1,c)=a$, obtemos $\xi'(a_1)=-[\partial_2 f(a)]^{-1}\circ[\partial_1 f(a)]$.

Observação: A afirmação de que um certo conjunto $G\subset\mathbb{R}^{m+n}$ é o gráfico de uma aplicação $\xi\colon V\to\mathbb{R}^n$ pressupõe uma decomposição em soma direta $\mathbb{R}^{m+n}=\mathbb{R}^m\oplus\mathbb{R}^n$. Ela significa que $G=\{(x,y)\in\mathbb{R}^{m+n}; x\in V,\ y=\xi(x)\}$. Todas as vezes que uma tal afirmação for feita, a decomposição $\mathbb{R}^{m+n}=\mathbb{R}^m\oplus\mathbb{R}^n$ deve estar fixada.

12 O teorema do posto

O *posto* de uma transformação linear $T\colon\mathbb{R}^m\to\mathbb{R}^n$ é a dimensão da sua imagem $T\cdot\mathbb{R}^m$, isto é, o número máximo de vetores linearmente independentes entre $T\cdot e_1,\dots,T\cdot e_m$, ou, equivalentemente, o número máximo de colunas linearmente independentes da matriz de T. Prova-se em Álgebra Linear que este é também o número máximo de linhas linearmente independentes dessa matriz e que o posto de T é igual a r se, e somente se, a matriz de T contém um determinante menor $r\times r$ não-nulo, mas qualquer determinante menor de ordem $r+1$ é igual a zero.

O *posto de uma aplicação diferenciável* $f\colon U\to\mathbb{R}^n$ num ponto $x\in U$ é o posto da sua derivada $f'(x)\colon\mathbb{R}^m\to\mathbb{R}^n$. Esse posto, evidentemente, não pode ser maior do que m nem maior do que n. Uma imersão $f\colon U\to$

$\mathbb{R}^n$, definida num aberto $U \subset \mathbb{R}^m$, tem posto m em todos os pontos $x \in U$ e uma submersão tem posto n em qualquer ponto. Por isso, imersões e submersões são chamadas *aplicações de posto máximo.*

Exemplo 23. O posto de uma aplicação diferenciável $f \colon U \to \mathbb{R}^n$, em geral, varia de ponto para ponto. Quando $n = 1$, o posto de f é 1 nos pontos regulares e zero nos pontos críticos de x. Se $f \colon U \to \mathbb{R}^2$ é holomorfa, digamos $f = u + iv$, então seu posto em cada ponto $z \in U$ só pode ser igual a 0 ou 2. Com efeito, sua derivada $f'(x) \colon \mathbb{R}^2 \to \mathbb{R}^2$ é a multiplicação por um número complexo, logo é igual a zero ou é uma transformação linear invertível (se tal número for $\neq 0$). Outra maneira de ver isto é olhar para a matriz jacobiana de f. Pela equação de Cauchy-Riemann

$$Jf = \begin{pmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y} \\ \\ -\dfrac{\partial u}{\partial y} & \dfrac{\partial u}{\partial x} \end{pmatrix}$$

logo $\det Jf = \left(\dfrac{\partial u}{\partial y}\right)^2 + \left(\dfrac{\partial u}{\partial x}\right)^2$ só se anula quando a matriz Jf for também nula. Finalmente, a aplicação $f \colon \mathbb{R}^2 \to \mathbb{R}^2$, dada por $f(x,y) = (x^3, y^2)$, tem matriz jacobina $Jf = \begin{pmatrix} 3x^2 & 0 \\ 0 & 2y \end{pmatrix}$. Se $x \neq 0$ e $y \neq 0$ então $Jf(x,y)$ é invertível e tem posto 2 no ponto (x,y). Nos pontos $(x,0)$ com $x \neq 0$ e nos pontos $(0,y)$ com $y \neq 0$, o posto de f é 1. Na origem $(0,0)$ o posto de f é zero.

Se $f \colon U \to \mathbb{R}^n$ é de classe C^1, o posto de f é uma "função semi-contínua inferiormente" (com valores inteiros). Isto significa que se f tem posto r num ponto $x \in U$ então existe uma bola B de centro a (contida em U) tal que em todo ponto de B o posto de f é $\geq r$. Com efeito, existe um determinante menor $r \times r$ da matriz $Jf(x)$ que é diferente de zero. Por continuidade, este mesmo menor é não-nulo em todos os pontos de uma bola B de centro x. Nesses pontos, o posto de f é, portanto, pelo menos igual a r.

Antes de demonstrarmos o Teorema do Posto, precisamos registrar dois fatos simples. O primeiro deles estende uma noção de convexidade introduzida no parágrafo 1 do Capítulo III.

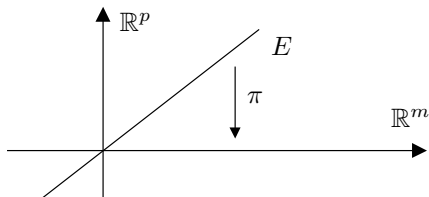
Dada uma decomposição $\mathbb{R}^{m+n} = \mathbb{R}^m \oplus \mathbb{R}^n$, diremos que um conjunto $X \subset \mathbb{R}^{m+n}$ é *verticalmente convexo* quando todo segmento vertical

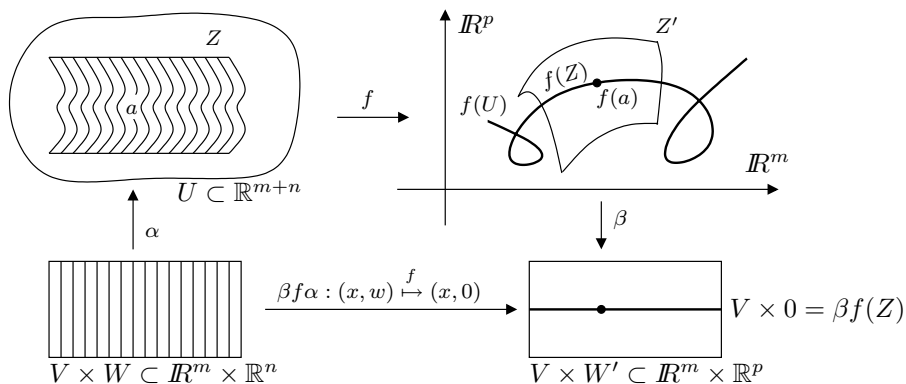
$[(x, y'), (x, y'')]$, cujas extremidades pertencem a X, está contido em X. Por exemplo, se $X = V \times W$ onde $V \subset \mathbb{R}^m$ é qualquer e $W \subset \mathbb{R}^n$ é convexo então X é verticalmente convexo.

Lema 1. *Seja $U \subset \mathbb{R}^{m+n} = \mathbb{R}^m \oplus \mathbb{R}^n$ um aberto verticalmente convexo. Se $f: U \to \mathbb{R}^p$ possui segunda derivada parcial $\partial_2 f$, a qual é identicamente nula em U, então f é independente da segunda variável, isto é, $f(x, y_1) = f(x, y_2)$ sejam quais forem (x, y_1) e (x, y_2) em U.*

Demonstração: Dados (x, y_1) e (x, y_2) em U, definamos o caminho $\lambda: [0, 1] \to \mathbb{R}^p$ pondo $\lambda(t) = f(x, (1-t)y_1 + ty_2)$. Para todo $t \in [0, 1]$, temos $\lambda'(t) = \partial_2 f(x, (1-t)y_1 + ty_2) \cdot (y_2 - y_1) = 0$. Logo λ é constante. Em particular, $\lambda(0) = \lambda(1)$, ou seja, $f(x, y_1) = f(x, y_2)$.

O segundo item do nosso registro é um resultado elementar de Álgebra Linear, que constitui o

Lema 2. *Seja $E \subset \mathbb{R}^{m+p}$ um subespaço vetorial de dimensão m. Existe uma decomposição em soma direta $\mathbb{R}^{m+p} = \mathbb{R}^m \oplus \mathbb{R}^p$ tal que a primeira projeção $\pi: \mathbb{R}^{m+p} \to \mathbb{R}^m$, $\pi(x, y) = x$, aplica E isomorficamente sobre $\mathbb{R}^m$.*

Demonstração: Seja A uma matriz $(m+p) \times m$ cujas colunas formam uma base de E. O posto de A é m, logo m de suas linhas, de índices $i_1 < \cdots < i_m$, são linearmente independentes. Estes índices determinam uma projeção $\pi: \mathbb{R}^{m+p} \to \mathbb{R}^m$, a qual transforma as colunas de A em m vetores linearmente independentes, logo aplica E isomorficamente sobre $\mathbb{R}^m$.

Teorema do Posto. *Seja $f: U \to \mathbb{R}^{m+p}$ de classe C^k $(k \geq 1)$ e posto constante m em cada ponto do aberto $U \subset \mathbb{R}^{m+n}$. Para cada ponto $a \in U$, existem difeomorfismos α, de um aberto em $\mathbb{R}^m \times \mathbb{R}^n$ sobre uma vizinhança aberta de a, e β, de uma vizinhança aberta de $f(a)$ sobre um aberto em $\mathbb{R}^m \times \mathbb{R}^p$, ambos, α e β, de classe C^k, tais que $\beta \circ f \circ \alpha: (x, y) \mapsto (x, 0)$.*

Descrição pictória do Teorema do Posto. Cada uma das fibras da vizinhança $Z \ni a$ é transformada por f num único ponto, do mesmo modo que cada segmento vertical $x \times W$ em $V \times W$ é transformado por $\beta f \alpha$ no ponto $(x, 0)$. O significado da frase, "do mesmo modo"é dado pelos difeomorfismos β e α. A restriçao $f \colon Z \to Z'$ é equivalente a $\beta f \alpha \colon V \times W \to V \times W'$.

Demonstração: Seja $E = f'(a) \cdot \mathbb{R}^{m+n} \subset \mathbb{R}^{m+p}$. Pelo Lema 2, como $\dim E = m$, existe uma decomposição em soma direta $\mathbb{R}^{m+p} = \mathbb{R}^m \oplus \mathbb{R}^p$ tal que a primeira projeção $\pi \colon \mathbb{R}^{m+p} \to \mathbb{R}^m$ aplica E isomorficamente sobre $\mathbb{R}^m$. Então $(\pi \circ f)'(a) = \pi \circ f'(a) \colon \mathbb{R}^{m+n} \to \mathbb{R}^m$ é sobrejetiva. Pela Forma Local das Submersões, existe um difeomorfismo $\alpha \in C^k$ de um aberto $V_0 \times W \subset \mathbb{R}^m \times \mathbb{R}^n$ sobre uma vizinhança aberta de a, tal que $\pi f \alpha(x, y) = x$. Isto significa que $f\alpha(x, y) = (x, \lambda(x, y))$, onde $\lambda \colon V_0 \times W \to \mathbb{R}^p$ é de classe C^k. Afirmamos que $\partial_2 \lambda = 0$. Com efeito, a matriz jacobiana de $f\alpha$ tem a forma $\begin{pmatrix} I & 0 \\ A & B \end{pmatrix}$, onde I é a matriz identidade $m \times m$ e $B = \partial\lambda_i / \partial\lambda_j$. Como o posto de $f\alpha$ (igual ao de f) é m, temos $B = 0$, donde $\partial_2 \lambda = 0$. Como W pode ser tomado convexo, segue-se do Lema 1 que $\lambda(x, y)$ não depende de y. Seja $\alpha(a_1, a_2) = a$. Considerando a injeção $i \colon V_0 \to V_0 \times W$, dada por $i(x) = (x, a_2)$, obtemos a aplicação de classe C^k, $f\alpha i \colon x \mapsto (x, \lambda(x, a_2))$, que tem derivada injetiva no ponto a_1 e cumpre $f\alpha i(x) = f\alpha(x, y)$ para qualquer $(x, y) \in V_0 \times W$. Pela Forma Local das Imersões, existe um difeomorfismo β de classe C^k numa vizinhança aberta de $f(a)$ sobre um aberto em $\mathbb{R}^m \times \mathbb{R}^p$, tal que $\beta f \alpha i(x) = (x, 0)$. Aqui, x está possivelmente num aberto $V \subset V_0$, contendo a_1. Como $f\alpha i(x) = f\alpha(x, y)$, temos $\beta f \alpha(x, y) = (x, 0)$ quaisquer que sejam $x \in V$, $y \in W$.

Corolário. *Seja* $f: U \to \mathbb{R}^n$ *de classe* C^1, *com posto constante no aberto* $U \subset \mathbb{R}^m$. *Então:*

1º) f *é localmente injetiva se, e somente se, é uma imersão;*

2º) f *é aberta se, e somente se, é uma submersão.*

Para provar o item 1º), observe que se o posto de f fosse $p < m$ então, no Teorema do Posto, a aplicação $\beta f \alpha: (x, y) \mapsto (x, 0)$, definida no produto $V \times W$ dos abertos $V \subset \mathbb{R}^p$ e $W \subset \mathbb{R}^{m-p}$, não seria injetiva, logo f não seria localmente injetiva. Assim, f localmente injetiva $\Rightarrow$ posto de $f = m \Rightarrow f$ é imersão. A recíproca é óbvia.

Quanto ao item 2º), como toda submersão de classe C^1 é aberta, basta provar que f aberta $\Rightarrow$ posto de $f = n$. Se $p =$ posto de f fosse menor do que n, dado o aberto $Z = \alpha(V \times W)$, com $\beta f \alpha: V \times W \to \mathbb{R}^p \times \mathbb{R}^{n-p}$ da forma $(x, y) \mapsto (x, 0)$, como no Teorema do Posto, a imagem $\beta f(Z) = \beta f \alpha(V \times W) = V \times \{0\}$ não seria aberta em $\mathbb{R}^n = \mathbb{R}^p \times \mathbb{R}^{n-p}$, logo $f(Z) \subset \mathbb{R}^n$ não seria um aberto.

O teorema seguinte serve de auxiliar ao uso do Teorema do Posto, mostrando que há sempre regiões onde uma aplicação de classe C^1 tem posto constante.

Teorema 7. *Seja* $f: U \to \mathbb{R}^n$ *de classe* C^1 *no aberto* $U \subset \mathbb{R}^m$. *Para cada* $r = 0, 1, \ldots, p = \min\{m, n\}$, *seja* A_r *o interior do conjunto dos pontos de* U *nos quais* f *tem posto* r. *O conjunto aberto* $A = A_0 \cup A_1 \cup \cdots \cup A_p$ *é denso em* U.

Demonstração: Devemos mostrar que, para cada aberto não vazio $V \subset U$, tem-se $V \cap A \neq \varnothing$. Como o posto de f só assume um número finito de valores ($\leq p$), existe um ponto $a \in V$ tal que em nenhum outro ponto de V o posto de f é maior do que $r =$ posto de $f'(a)$. Podemos tomar uma bola aberta B, de centro a, tal que o posto de f em todos os pontos de B é $\geq r$. Escolhendo B suficientemente pequena, teremos $B \subset V$ e então o posto de f em todos os pontos de B será $\geq r$, donde exatamente $= r$. Assim $B \subset A_r$, donde $V \cap A \supset B \neq \varnothing$.

Observação: Em geral, vários dos conjuntos A_r do teorema acima são vazios. Note que A_p (que é igual a A_m se $m \leq n$ e é igual a A_n quando $m \geq n$) é o conjunto dos pontos $x \in U$ nos quais o posto de $f'(x)$ é igual a p; tal conjunto é sempre aberto logo, no caso $r = p$, não é preciso tomar o interior.

Corolário 1. *Seja* $f: U \to \mathbb{R}$ *de classe* C^1. *Existe um subconjunto aberto e denso* $A \subset U$ *tal que* f *tem posto constante em cada componente*

conexa de A.

Com efeito, como em $A = A_0 \cup A_1 \cup \cdots \cup A_p$ os abertos A_r são dois a dois disjuntos, cada componente conexa de A está contida em algum A_r, logo o posto de f é constante nela.

Corolário 2. *Se a aplicação $f\colon U \to \mathbb{R}^n$, de classe C^1 no aberto $U \subset \mathbb{R}^m$, é localmente injetiva, então $m \leq n$ e o conjunto dos pontos $x \in U$ nos quais $f'(x)\colon \mathbb{R}^m \to \mathbb{R}^n$ é injetiva é aberto e denso em U.*

Com efeito, considerando a decomposição $A = A_0 \cup \cdots \cup A_p$, dada pelo Teorema 7, em cada $A_r \neq \varnothing$ f é localmente injetiva com posto constante. Pelo Corolário do Teorema do Posto, $f|A_r$ é uma imersão. Ou seja, $A_r \neq \varnothing \Rightarrow r = m$. Assim, $A = A_m$, donde $m \leq n$ e $f'(x)$ é injetiva para todo x pertencente ao aberto denso $A \subset U$.

Corolário 3. *Se a aplicação $f\colon U \to \mathbb{R}^n$, de classe C^1 no aberto $U \subset \mathbb{R}^m$ é aberta então $m \geq n$ e o conjunto dos pontos $x \in U$ nos quais a derivada $f'(x)\colon \mathbb{R}^m \to \mathbb{R}^n$ é sobrejetiva é aberto e denso em U.*

Demonstração análoga a do corolário anterior.

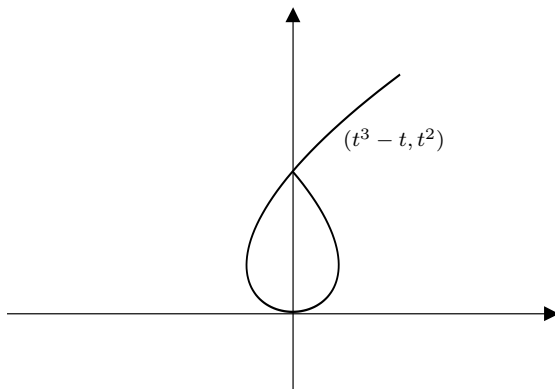
Observação: Quando $m = 1$, dizer que $f'(x)$ é injetiva equivale a afirmar que o caminho $f\colon I \to \mathbb{R}^n$ tem vetor-velocidade $\neq 0$ no ponto $x \in I$. Neste caso, o Corolário 2 não depende do Teorema do Posto: o conjunto $A = \{x \in I; f'(x) \neq 0\}$ é sempre aberto. Se não fosse denso em I, seu complementar conteria um intervalo não-degenerado $J \subset I$. Em cada ponto $y \in J$ seria $f'(y) = 0$, logo f seria constante em J, e assim perderia sua injetividade local. Analogamente, para $n = 1$, o Corolário 3 pode ser demonstrado sem ajuda do Teorema do Posto. neste caso, dizer que a função $f\colon U \to \mathbb{R}$ tem derivada sobrejetiva no ponto $x \in U$ significa afirmar que $df(x) \neq 0$. O conjunto $A = \{x \in U; df(x) \neq 0\}$ é sempre aberto. Se não fosse denso em U, seu complementar conteria uma bola aberta B. Em cada ponto $y \in B$, seria $df(y) = 0$, logo f seria constante portanto $f(B) =$ ponto e f não seria uma aplicação aberta.

13 Superfícies no espaço euclidiano

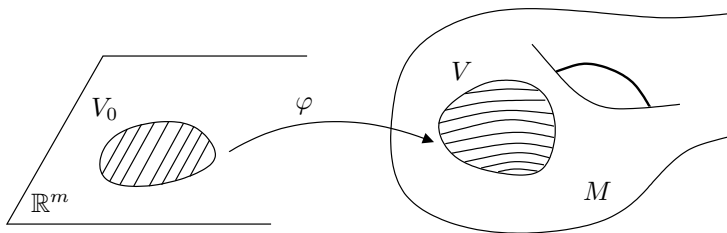
O Teorema da Aplicação Inversa e suas conseqüêncis estabelecidas nos parágrafos anteriores encontram naturalmente aplicações na teoria das superfícies, da qual faremos aqui um esboço introdutório.

Uma *parametrização de classe* C^k de um conjunto $V \subset \mathbb{R}^n$ é um homeomorfismo $\varphi\colon V_0 \to V$, que é também uma imersão de classe C^k, definida no aberto $V_0 \subset \mathbb{R}^m$. Tem-se necessariamente $m \leq n$. Quando for preciso ser explicito, dir-se-á que φ é uma *parametrização* m*-dimensional* do conjunto V.

A definição acima, mencionando "imersão", só faz sentido quando $k \geq 1$. Às vezes é interessante considerar *parametrizações de classe* C^0, que são apenas homeomorfismos $\varphi\colon V_0 \to V$, de um aberto $V_0 \subset \mathbb{R}^m$ sobre um conjunto $V \subset \mathbb{R}^n$.

Exemplo 24. Uma imersão injetiva $\varphi\colon V_0 \to \mathbb{R}^n$, mesmo de classe C^∞, pode não ser um homeomorfismo de V_0 sobre $V = f(V_0)$. Um exemplo disso é o caminho $f\colon (-1, +\infty) \to \mathbb{R}^2$, definido por $f(t) = (t^3 - t, t^2)$. Vê-se facilmente que f é uma bijeção do aberto $V_0 = (-1, +\infty) \subset \mathbb{R}$ sobre $V = f(V_0) \subset \mathbb{R}^2$, que $f \in C^\infty$ e que $f'(t) = (3t^2 - 1, 2t)$ nunca se anula. Mas se tomarmos uma seqüência decrescente de números reais t_n, com $\lim t_n = -1$, veremos que $\lim\limits_{n\to\infty} f(t_n) = (0, 1) = f(1)$ sem que seja $\lim t_n = 1$. Portanto $f^{-1}\colon V \to V_0$ não é contínua e f não é um homeomorfismo de V_0 sobre V.

Se $\varphi\colon V_0 \to V$ é uma parametrização, cada ponto $y = \varphi(x) \in V$, embora pertença a $\mathbb{R}^n$, necessita apenas de m números para ter sua posição determinada, a saber, as m coordenadas de $x \in V_0$, chamadas os *parâmetros* que servem para localizar os pontos de V.

Todo sistema de coordenadas num aberto $U \subset \mathbb{R}^n$ é uma parametrização de U. (Veja §9.) Quando $m < n$, a imagem V de uma parametrização $\varphi\colon V_0 \to V$, definida em $V_0 \subset \mathbb{R}^m$, é um conjunto com interior vazio em $\mathbb{R}^n$, em virtude da Forma Local das Imersões.

Uma *superfície* de dimensão m e classe C^k em $\mathbb{R}^n$ é um conjunto $M \subset \mathbb{R}^n$ que pode ser coberto por uma coleção de abertos $U \subset \mathbb{R}^n$, tais que cada $V = U \cap M$ admite uma parametrização $\varphi\colon V_0 \to V$, de classe C^k, definida num aberto $V_0 \subset \mathbb{R}^m$. Cada

um desses conjuntos $V = U \cap M$ é um aberto em M. Para cada $p \in M$, diz-se que V é uma *vizinhança parametrizada* de p.

Assim, uma superfície de classe C^k e dimensão m em $\mathbb{R}^n$ é um subconjunto tal que cada um dos seus pontos possui uma vizinhança parametrizada, por uma parametrização de classe C^k e dimensão m.

Quando $M \subset \mathbb{R}^n$ é uma superfície de dimensão m, escrevemos às vezes M^m. Dizemos também que M tem *codimensão* $n-m$. Se $\dim M = 1$, M chama-se uma *curva* de classe C^k.

Como a restrição de uma parametrização $\varphi\colon V_0 \to V$ a um aberto $W_0 \subset V_0$ é ainda uma parametrização de $W = \varphi(W_0)$, todo ponto p de uma superfície M possui vizinhanças parametrizadas arbitrariamente pequenas: se $\varphi\colon V_0 \to V$ é uma parametrização de $V = U \cap M$ e B é qualquer bola aberta de centro p contida em U então $W = B \cap M$ é parametrizada pela restrição de φ a $W_0 = \varphi^{-1}(W)$.

Exemplo 25. Todo aberto $U \subset \mathbb{R}^n$ é uma superfície de dimensão n e classe C^∞ em $\mathbb{R}^n$, coberta por uma única vizinhança parametrizada U, sendo $\varphi\colon U \to U$ a aplicação identidade. Reciprocamente, se $M \subset \mathbb{R}^n$ é uma superfície de dimensão n e classe C^1 então cada vizinhança parametrizada $V \subset M$, sendo imagem do aberto $V_0 \subset \mathbb{R}^n$ pelo difeomorfismo $\varphi\colon V_0 \to V$, é um subconjunto aberto de $\mathbb{R}^n$, pelo Teorema da Aplicação Inversa. O conjunto M, reunião das vizinhanças parametrizadas V, é portanto aberto em $\mathbb{R}^n$. Assim, as superfícies de classe C^1 e dimensão n em $\mathbb{R}^n$ são os subconjuntos abertos de $\mathbb{R}^n$. [Isto também é verdade para classe C^0, e decorre do Teorema da Invariância do Domínio, de Brouwer.]

Exemplo 26. No outro extremo, $M \subset \mathbb{R}^n$ é uma superfície de dimensão 0 se, e somente se, para cada ponto $p \in M$ existe um aberto $U \subset \mathbb{R}^n$ tal que $U \cap M$ admite uma parametrização $\varphi \colon V_0 \to V$, definida num aberto $V_0 \subset \mathbb{R}^0 = \{0\}$. Assim, V_0 é um ponto e $V = \{p\}$. Portanto, $M \subset \mathbb{R}^n$ é uma superfície de dimensão 0 se, e somente se, é um conjunto discreto (isto é, seus pontos são todos isolados).

Exemplo 27. Sejam $\mathbb{R}^{m+n} = \mathbb{R}^m \oplus \mathbb{R}^n$ uma decomposição em soma direta e $f \colon U \to \mathbb{R}^n$ uma aplicação de classe C^k, definida no aberto $U \subset \mathbb{R}^m$. O conjunto $M = \{(x, y) \in \mathbb{R}^{m+n}; x \in U, y = f(x)\}$, *gráfico* da aplicação f, é uma superfície de dimensão m e classe C^k em $\mathbb{R}^{m+n}$, imagem de uma única parametrização $\varphi \colon U \to M$, $\varphi(x) = (x, f(x))$. A respeito deste exemplo, observamos o seguinte: M só será uma superfície de classe C^{k+1} se a aplicação f também for de classe C^{k+1}. Com efeito, consideremos a projeção $\pi \colon \mathbb{R}^{m+n} \to \mathbb{R}^m$. Sabemos, desde o Capítulo I, que a restrição $\pi|M$ é um homeomorfismo de M sobre U. Sabemos também que $\pi \in C^\infty$. Dado um ponto arbitrário $a \in U$, suponhamos que existam uma vizinhança V de $\varphi(a)$ em M e $\xi \colon V_0 \to V$, uma parametrização de classe C^{k+1}. Então $\pi \circ \xi = (\pi|M) \circ \xi$ é um homeomorfismo de classe C^{k+1} de V_0 sobre o aberto $\pi(V) = W_0 \subset U$. Para todo ponto $x \in V_0$, a Regra da Cadeia aplicada à igualdadde $\varphi \circ (\pi \circ \xi) = \xi$ dá $\varphi'(\pi\xi(x)) \circ (\pi \circ \xi)'(x) = \xi'(x)$. Como a transformação linear $\xi'(x)$ é injetiva, seja qual for $x \in V_0$, segue-se daí que, para todo $x \in V_0$, $(\pi \circ \xi)'(x) \colon \mathbb{R}^m \to \mathbb{R}^m$ é injetiva e portanto um isomorfismo. Pelo teorema da diferenciabilidade do homeomorfismo inverso, $\pi \circ \xi \colon V_0 \to W_0$ é um difeomorfismo de classe C^{k+1}. Então $\varphi|W_0 = \xi \circ (\pi \circ \xi)^{-1} \colon W_0 \to V$ é uma parametrização de C^{k+1} de V. Em particular, φ é de classe C^{k+1} no aberto W_0 com $a \in W_0 \subset U$. Como a é qualquer ponto de U, vemos que $\varphi \colon x \mapsto (x, f(x))$ é uma aplicação de classe C^{k+1} em U, logo $f \colon U \to \mathbb{R}^n$ é de classe C^{k+1}. Portanto, se tomarmos uma aplicação que seja de classe C^k, mas não de classe C^{k+1}, seu gráfico fornecerá um exemplo de superfície de classe C^k mas não C^{k+1}. Por exemplo, seja $f \colon \mathbb{R} \to \mathbb{R}$, $f(x) = x^{4/3}$. O gráfico de f é uma curva de classe C^1 em $\mathbb{R}^2$ mas não é de classe C^2.

Exemplo 28. Ser uma superfície é uma propriedade local, isto é, se cada ponto p do conjunto M pertence a um aberto $Z \subset \mathbb{R}^n$ tal que $Z \cap M$ é uma superfície de dimensão m e classe C^k, então M é uma superfície de dimensão m e classe C^k. Assim, por exemplo, se para todo ponto $p \in M$ existe um aberto $Z \subset \mathbb{R}^n$, tal que $p \in Z$ e $Z \cap M$ é o gráfico

de uma aplicação $f\colon V_0 \to \mathbb{R}^n$, de classe C^k num aberto $V_0 \subset \mathbb{R}^m$, então M é uma superfície de dimensão m e classe C^k. Em particular, toda *hiperfície* de classe C^k em $\mathbb{R}^{n+1}$ é uma superfície de dimensão n e classe C^k. (Veja §9 do Capítulo III.)

Exemplo 29. Se $M_1 \subset \mathbb{R}^{n_1}$ e $M_2 \subset \mathbb{R}^{n_2}$ são superfícies de classe C^k e dimensões m_1, m_2 respectivamente, o *produto cartesiano* $M_1 \times M_2 \subset \mathbb{R}^{n_1+n_2}$ é uma superfície de classe C^k e dimensão m_1+m_2: se $\varphi\colon V_0 \to V$ e $\psi\colon W_0 \to W$ são respectivamente parametrizações em M_1 e M_2 então $\varphi \times \psi\colon V_0 \times W_0 \to V \times W$, definida por $(\varphi \times \psi)(x, y) = (\varphi(x), \psi(y))$ é uma parametrização em $M_1 \times M_2$. Evidentemente, o produto cartesiano de um número finito qualquer de superfícies é ainda uma superfície. Em particular, como o círculo $S^1 \subset \mathbb{R}^2$ é uma curva ("superfície"de dimensão 1) de classe C^∞, o *toro bidimensional* $T^2 = S^1 \times S^1 \subset \mathbb{R}^4$ é uma superfície C^∞. Para todo n, o *toro n-dimensional* $T^n = S^1 \times \cdots \times S^1 \subset \mathbb{R}^{2n}$ é uma superfície C^∞ de dimensão n em $\mathbb{R}^{2n}$.

Dada uma parametrização $\varphi\colon V_0 \to V$, de classe C^k $(k \geq 1)$, do ponto de vista em que nos colocamos até agora não tem sentido indagar se o homeomorfismo $\varphi^{-1}\colon V \to V_0$ é uma aplicação diferenciável, já que seu domínio não é necessariamente um subconjunto aberto do espaço euclidiano. Assim sendo, não podemos simplesmente usar a Regra da Cadeia para justificar a validez do teorema seguinte.

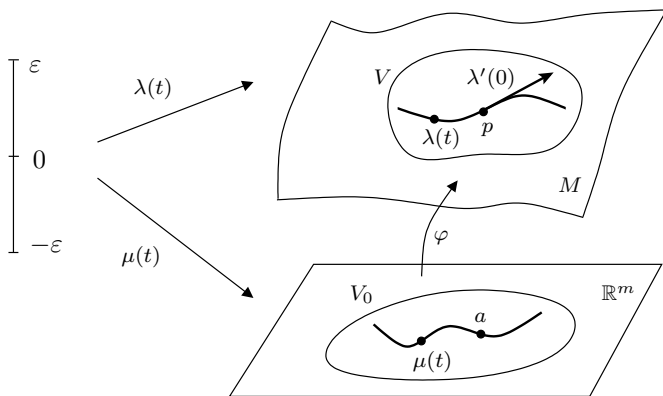
Teorema 9. *Sejam $M \subset \mathbb{R}^n$ uma superfície de classe C^k e dimensão m e $f\colon U \to \mathbb{R}^n$ uma aplicação de classe C^k (respectivamente diferenciável no ponto $a \in U$), definida no aberto $U \subset \mathbb{R}^p$. Se $f(U) \subset V$, onde V é a imagem de uma parametrização $\varphi\colon V_0 \to V \subset M$ (de classe C^k) então a aplicação composta $\varphi^{-1} \circ f\colon U \to \mathbb{R}^m$ é de classe C^k (respectivamente diferenciável no ponto a).*

Demonstração: Dado qualquer $a \in U$, seja $x_0 \in V_0$ tal que $\varphi(x_0) = f(a)$. Pelo Corolário da Forma Local das Imersões, existe W aberto, com $x_0 \in W \subset V_0$, tal que $\varphi^{-1}\colon \varphi(W) \to W$ é a restrição de uma aplicação $\xi\colon Z \to \mathbb{R}^m$, de classe C^k no aberto Z, com $\varphi(W) \subset Z \subset \mathbb{R}^n$. Como $\varphi\colon V_0 \to V$ é homeomorfismo, $\varphi(W)$ é um subconjunto aberto de V. Logo existe U_1 aberto, com $a \in U_1 \subset U$ e $f(U_1) \subset \varphi(W)$. Assim, para cada $x \in U_1$, $\varphi^{-1}(f(x)) = \xi(f(x))$, isto é, $\varphi^{-1} \circ f = \xi \circ f$ em U_1. Como $\xi \in C^k$, isto prova o Teorema.

Vejamos agora algumas conseqüências do Teorema 9. A primeira é a definição do espaço vetorial tangente a uma superfície num dos seus

pontos.

Seja $M \subset \mathbb{R}^n$ uma superfície de dimensão m e classe C^k $(k \geq 1)$. Dado um ponto $p \in M$, o *espaço vetorial tangente a* M no ponto p é o subespaço vetorial $T_pM \subset \mathbb{R}^n$ que pode ser descrito de duas maneiras:
A) Tomamos uma parametrização $\varphi\colon V_0 \to V$, de classe C^k, de uma vizinhança $V \ni p$; seja $\varphi(a) = p$. Pomos $T_pM = \varphi'(a) \cdot \mathbb{R}^m$ = imagem da transformação linear injetiva $\varphi'(a)\colon \mathbb{R}^m \to \mathbb{R}^n$;
B) Consideramos todos os caminhos $\lambda\colon (-\varepsilon, \varepsilon) \to M$, com $\lambda(0) = p$, diferenciáveis no ponto 0. Definimos T_pM como o conjunto dos vetores-velocidade $\lambda'(0)$ desses caminhos.

A definição A) apresenta T_pM como um subespaço vetorial de dimensão m do espaço $\mathbb{R}^n$ mas tem o inconveniente de depender de uma parametrização φ. [Não está claro que, se tomássemos outra parametrização ψ, com $\psi(b) = p$, teríamos $\psi'(b) \cdot \mathbb{R}^m = \varphi'(a) \cdot \mathbb{R}^m$.] Por outro lado, a definição B) não depende da escolha de parametrização mas não é evidente que T_pM, assim definido, seja um espaço vetorial. Ambos os inconvenientes serão sanados se provarmos que A) e B) definem o mesmo conjunto. É isto que faremos agora.

Em primeiro lugar, todo vetor $w \in \varphi'(a) \cdot \mathbb{R}^m$, sendo da forma $w = \varphi'(a) \cdot v = \dfrac{\partial \varphi}{\partial v}(a)$, é o vetor-velocidade, no ponto $t = 0$, do caminho λ, dado por $\lambda(t) = \varphi(a + tv)$. Reciprocamente, se $\lambda\colon (-\varepsilon, \varepsilon) \to M$ é um caminho diferenciável no ponto 0, com $\lambda(0) = p$ podemos, sem perda de generalidade, supor $\varepsilon > 0$ tão pequeno que a imagem de λ esteja contida na vizinhança V. Então, pelo Teorema 9, o caminho $\mu = \varphi^{-1} \circ \lambda\colon (-\varepsilon, \varepsilon) \to \mathbb{R}^m$ é diferenciável no ponto 0. Como $\varphi \circ \mu = \lambda$, a Regra da Cadeia nos dá $\lambda'(0) = \varphi'(a) \cdot \mu'(0)$, logo $\lambda'(0) \in \varphi'(a) \cdot \mathbb{R}^m$.

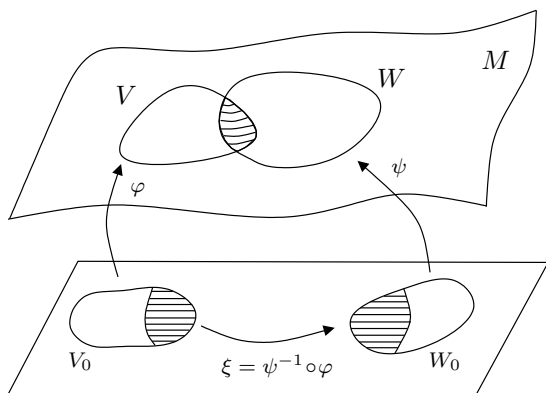
A toda parametrização $\varphi\colon V_0 \to V$, com $p = \varphi(a) \in V$, corresponde uma base do espaço vetorial T_pM, formada pelos vetores

$$\frac{\partial\varphi}{\partial x_1}(a) = \varphi'(a)\cdot e_1, \ldots, \frac{\partial\varphi}{\partial x_m}(a) = \varphi'(a)\cdot e_m\,.$$

Exemplo 30. Se $M \subset \mathbb{R}^{m+n}$ é o gráfico de uma aplicação $f\colon U \to \mathbb{R}^n$, seu espaço vetorial tangente T_pM, no ponto $p = (a, f(a))$, é o gráfico da derivada $f'(a)\colon \mathbb{R}^m \to \mathbb{R}^n$, ou seja $T_pM = \{(v, f'(a)\cdot v); v \in \mathbb{R}^m\}$. Basta tomar a parametrização natural $\varphi\colon U \to M$, $\varphi(x) = (x, f(x))$, cuja derivada no ponto a é dada por $\varphi'(a)\cdot v = (v, f'(a)\cdot v)$, para ver que $T_pM = \varphi'(a)\cdot\mathbb{R}^m$ é o gráfico de $f'(a)$. A mesma observação se aplica a uma superfície que é localmente o gráfico de uma aplicação.

A segunda conseqüência do Teorema 9 se refere à "mudança de parametrizações". Vamos enunciá-la como o

Teorema 10. *Seja $\psi\colon W_0 \to V \subset \mathbb{R}^n$ uma parametrização m-dimensional de classe C^k. A fim de que uma aplicação $\varphi\colon V_0 \to V$, de classe C^k, seja uma parametrização de V é necessário e suficiente que $\varphi = \psi\circ\xi$, onde $\xi\colon V_0 \to W_0$ é um difeomorfismo C^k. [Em particular, φ deve também ser m-dimensional.]*

Demonstração: A suficiência é óbvia. Quanto à necessidade, se φ é uma parametrização, pelo Teorema 9 as aplicações $\xi = \psi^{-1}\circ\varphi\colon V_0 \to W_0$ e $\xi^{-1} = \varphi^{-1}\circ\psi\colon W_0 \to V_0$ são ambas de classe C^k, logo são difeomorfismos C^k, com $\varphi = \psi\circ\xi$.

A utilização mais freqüente do Teorema 10 se faz assim: sejam V, W vizinhanças parametrizadas numa superfície M. Temos as parametrizações $\varphi\colon V_0 \to V$ e $\psi\colon W_0 \to W$. Se $V \cap W \neq \varnothing$ então cada ponto

$p \in V \cap W$ pode ser caracterizado pelos parâmetros $x_1, \ldots, x_m$, coordenadas do ponto x tal que $\varphi(x) = p$, ou pelas coordenadas $y_1, \ldots, y_m$ do ponto y tal que $\psi(y) = p$. A mudança de parâmetros $x \mapsto y$ define um difeomorfismo $\xi = \psi^{-1} \circ \varphi \colon \varphi^{-1}(V \cap W) \to \psi^{-1}(V \cap W)$, $\xi(x) = y$.

Se $p = \varphi(a) = \psi(b) \in V \cap W$, temos as bases

$$\left\{ \frac{\partial \varphi}{\partial x_1}(a), \ldots, \frac{\partial \varphi}{\partial x_m}(a) \right\} \quad \text{e} \quad \left\{ \frac{\partial \psi}{\partial y_1}(b), \ldots, \frac{\partial \psi}{\partial y_m}(b) \right\}$$

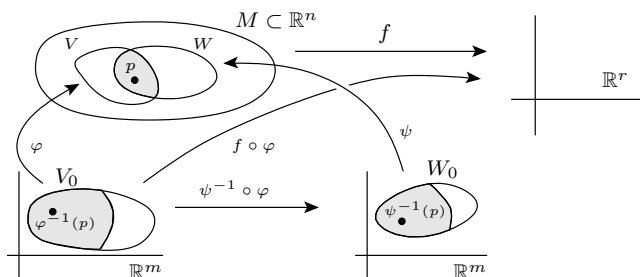
do espaço vetorial T_pM, determinadas pelas parametrizações $\varphi \colon V_0 \to V$ e $\psi \colon W_0 \to W$, respectivamente. Vale, evidentemente, uma relação do tipo $\frac{\partial \varphi}{\partial x_j}(a) = \sum_{i=1}^{m} \alpha_{ij} \frac{\partial \psi}{\partial y_i}(b)$. Para determinar os coeficientes α_{ij}, basta notar que, se $\xi_1, \ldots, \xi_m$ são as funções-coordenada de ξ, a igualdade $\varphi = \psi \circ \xi$ nos dá, pela Regra da Cadeia,

$$\frac{\partial \varphi}{\partial x_j}(a) = \sum_{i=1}^{m} \frac{\partial \psi}{\partial y_i}(b) \cdot \frac{\partial \xi_i}{\partial x_j}(a),$$

portanto $\alpha_{ij} = \frac{\partial \xi_i}{\partial x_j}(a)$. É costume indicarem-se também as funções-coordenada de ξ por $y_1, \ldots, y_m$ em vez de ξ_i. Neste caso, a mudança de base em T_pM correspondente à mudança de parametrização se exprime:

$$\frac{\partial \varphi}{\partial x_j}(a) = \sum_{i=1}^{m} \frac{\partial y_i}{\partial x_j}(a) \cdot \frac{\partial \psi}{\partial y_i}(b).$$

Qualquer que seja a notação empregada, o importante é notar que a matriz (α_{ij}), de passagem da base $\left(\frac{\partial \psi}{\partial y_i}(b)\right)$ para a base $\left(\frac{\partial \varphi}{\partial x_i}(a)\right)$, no espaço vetorial T_pM (onde $p = \varphi(a) = \psi(b)$) é a matriz jacobiana $J\xi(a)$, onde $\xi = \psi^{-1} \circ \varphi$.

A terceira conseqüência do Teorema 9 é a definição de aplicação diferenciável numa superfície. Antes de demonstrá-lo, dizíamos que a noção de aplicação diferenciável só tinha sentido quando o domínio era um aberto do espaço euclidiano. Agora podemos estender o conceito, de modo a abranger aplicações definidas em superfícies.

Seja $M \subset \mathbb{R}^n$ uma superfície de classe C^k $(k \geq 1)$. Uma aplicação $f\colon M \to \mathbb{R}^r$ diz-se *diferenciável* no ponto $p \in M$ quando existe uma parametrização C^k, $\varphi\colon V_0 \to V$, de uma vizinhança V de p em M, tal que $f \circ \varphi\colon V_0 \to \mathbb{R}^r$ (aplicação definida num aberto $V_0 \subset \mathbb{R}^m$) é diferenciável no ponto $a = \varphi^{-1}(p)$. Se tomarmos qualquer outra parametrização C^k, $\psi\colon W_0 \to W$, de uma vizinhança de p em M, teremos $f \circ \psi$ diferenciável no ponto $\psi^{-1}(p)$ se, e somente se, $f \circ \varphi$ for diferenciável no ponto $\varphi^{-1}(p)$, pois $f \circ \psi = (f \circ \varphi) \circ (\varphi^{-1} \circ \psi)$ e sabemos que $\varphi^{-1} \circ \psi\colon \psi^{-1}(V \cap W) \to \varphi^{-1}(V \cap W)$ é um difeomorfismo de classe C^k.

Analogamente, dada a aplicação $f\colon M \to \mathbb{R}^r$, definida numa superfície de classe C^k, dizemos que *f é de classe C^i* $(0 \leq i \leq k)$ quando, para cada ponto $p \in M$, existe uma parametrização $\varphi\colon V_0 \to V \subset M$, com $p \in V$, $\varphi \in C^k$, tal que $f \circ \varphi\colon V_0 \to \mathbb{R}^r$ é de classe C^i no aberto $V_0 \subset \mathbb{R}^m$. Neste caso, para qualquer outra parametrização $\psi\colon W_0 \to W \subset M$, com $p \in W$ e $\psi \in C^k$ tem-se ainda $f \circ \psi\colon W_0 \to \mathbb{R}^r$ de classe C^i. Mas se for $i \geq k+1$, não tem sentido definir a noção de aplicação de classe C^i em M se esta superfície é apenas de classe C^k e não mais.

Se $f\colon M \to \mathbb{R}^r$ é diferenciável no ponto $p \in M$, sua *derivada* nesse ponto é a transformação linear $f'(p)\colon T_pM \to \mathbb{R}^r$ definida do seguinte modo. Tomamos uma parametrização $\varphi\colon V_0 \to V$, com $p = \varphi(a)$. Dado um vetor $v \in T_pM$, temos $v = \varphi'(a) \cdot v_0$ para um vetor $v_0 \in \mathbb{R}^m$. Pomos $f'(p) \cdot v = (f \circ \varphi)'(a) \cdot v_0$.

Devemos mostrar que a transformação linear $f'(p)\colon T_pM \to \mathbb{R}^r$ está bem definida, ou seja, que se $\psi\colon W_0 \to W$ for outra parametrização, com $p = \psi(b)$, e tivermos $v = \psi'(b) \cdot w_0$ para algum, $w_0 \in \mathbb{R}^m$, então $(f \circ \varphi)'(a) \cdot v_0 = (f \circ \psi)'(b) \cdot w_0$. Com efeito, sabemos, pelo Teorema 10, que $\psi = \varphi \circ \zeta$, onde $\zeta\colon \psi^{-1}(V \cap W) \to \varphi^{-1}(V \cap W)$ é um difeomorfismo

de classe C^k entre abertos de $\mathbb{R}^n$, com $\zeta(b) = a$. Temos

$$\varphi'(a) \cdot v_0 = v = \psi'(b) \cdot w_0 = (\varphi \circ \zeta)'(b) \cdot w_0 = \varphi'(a) \cdot \zeta'(b) \cdot w_0 \,.$$

Como $\varphi'(a)$ é injetiva, segue-se que $\zeta'(b) \cdot w_0 = v_0$. Por conseguinte:

$$(f \circ \psi)'(b) \cdot w_0 = (f \circ \varphi \circ \zeta)'(b) \cdot w_0 = (f \circ \varphi)'(a) \cdot \zeta'(b) \cdot w_0 = \\ = (f \circ \varphi)'(a) \cdot v_0 \,,$$

como estávamos querendo.

Qualquer vetor $v \in T_pM$ é o vetor-velocidade, $v = \lambda'(0)$, de um caminho $\lambda\colon (-\varepsilon, \varepsilon) \to M$, com $\lambda(0) = p$. Então a imagem $f'(p) \cdot v = (f \circ \lambda)'(0)$ é o vetor-velocidade, no ponto 0, do caminho $f \circ \lambda\colon (-\varepsilon, \varepsilon) \to \mathbb{R}^r$.

Se $N \subset \mathbb{R}^r$ é outra superfície de classe C^k e a aplicação $f\colon M \to \mathbb{R}^r$, diferenciável no ponto $p \in M$, é tal que $f(M) \subset N$, diremos que $f\colon M \to N$ é uma *aplicação diferenciável* no ponto p. A observação que acabamos de fazer sobre $f'(b) \cdot v$ como vetor-velocidade de um caminho mostra que se $f\colon M \to N$ é diferenciável no ponto $p \in M$ então a derivada $f'(p)$ é uma transformação linear de T_pM em T_qN, onde $q = f(p)$.

Vale a Regra da Cadeia: se $f\colon M \to N$ é diferenciável no ponto $p \in M$ e $g\colon N \to \mathbb{R}^s$ é diferenciável no ponto $q = f(p)$, então $g \circ f\colon M \to \mathbb{R}^s$ é diferenciável no ponto p, com $(g \circ f)'(p) = g'(q) \cdot f'(p)$.

Generalizando a definição dada no Capítulo III, diremos que um ponto $c \in \mathbb{R}^n$ é *valor regular* de uma aplicação diferenciável $f\colon U \to \mathbb{R}^n$, definida num aberto $U \subset \mathbb{R}^m$, quando, para todo ponto $x \in U$ tal que $f(x) = c$, a derivada $f'(x)\colon \mathbb{R}^m \to \mathbb{R}^n$ for uma transformação linear sobrejetiva.

Considerando as funções-coordenada $f_1, \ldots, f_n\colon U \to \mathbb{R}$ da aplicação f, vemos que $c \in \mathbb{R}^n$ é um valor regular de f se, e somente se, em todo ponto $x \in f^{-1}(c)$ os vetores $\operatorname{grad} f_1(x), \ldots, \operatorname{grad} f_n(x)$ são linearmente independentes. Com efeito, estes vetores são as linhas da matriz jacobiana $Jf(x)$ e sua independência linear significa que esta matriz $n \times m$ tem posto n, ou seja, que a transformação linear $f'(x)$ é sobrejetiva. Quando $n = 1$, esta condição significa que $f(x) = c \Rightarrow \operatorname{grad} f(x) \neq 0$ e recaímos na definição de valor regular dada no Capítulo III. Como antes, observamos que a presente definição está formulada de tal modo que se nenhum $x \in U$ cumpre $f(x) = c$ (ou seja, se $f^{-1}(c) = \varnothing$) então c é valor regular de f.

Teorema 11. *Seja $f\colon U \to \mathbb{R}^{n-m}$ de classe C^k no aberto $U \subset \mathbb{R}^n$. Se $c \in \mathbb{R}^{n-m}$ é valor regular de f então $f^{-1}(c)$ é uma superfície de classe C^k e dimensão m em $\mathbb{R}^n$. Em cada ponto $p \in f^{-1}(c)$, o espaço tangente $T_p[f^{-1}(c)]$ é o núcleo da derivada $f'(p)\colon \mathbb{R}^n \to \mathbb{R}^{n-m}$.*

Demonstração: Pelo Teorema da Aplicação Implícita, para cada ponto $p \in f^{-1}(c)$ existe um aberto $Z \subset \mathbb{R}^n$, contendo p, tal que $Z \cap f^{-1}(c)$ é o gráfico de uma aplicação de classe C^k, definida num aberto de $\mathbb{R}^m$. Logo cada $Z \cap f^{-1}(c)$ é uma superfície de dimensão m e classe C^k. Segue-se que $f^{-1}(c)$ também o é. (Veja Exemplos 27 e 28.) Dado qualquer $p \in f^{-1}(c)$, como todo vetor v, tangente a $f^{-1}(c)$ no ponto p, é o vetor-velocidade de um caminho λ contido em $f^{-1}(c)$ [logo $f \circ \lambda =$ constante], temos $f'(p) \cdot v = (f \circ \lambda)'(0) = 0$. Assim o espaço vetorial tangente $T_p[f^{-1}(c)]$ está contido no núcleo de $f'(p)$. Como ambos são espaços vetoriais de dimensão m, segue-se que eles são iguais.

Observações: 1. Para o leitor que se interessa por detalhes desta natureza, chamamos a atenção para o fato de que, na definição de superfície, em nenhum lugar se exige a existência de algum ponto. Assim, estritamente falando, o conjunto vazio é uma superfície de classe C^∞ (e dimensão m, para todo $m \in \mathbb{N}$).

2. A imagem inversa, $f^{-1}(c)$, de um valor regular é um subconjunto fechado do domínio U de f mas não é necessariamente fechado em $\mathbb{R}^n$, muito menos conexo: tome, por exemplo, $U = \mathbb{R}^2 - \{0\}$ e $f\colon U \to \mathbb{R}$ dada por $f(x, y) = xy$. Então $df(x, y) = ydx + xdy$. Para todo $(x, y) \in U$, vale $df(x, y) \neq 0$, logo todo número real c é valor regular de f. Se $c \neq 0$ então $f^{-1}(c)$ é uma hipérbole, subconjunto fechado desconexo de $\mathbb{R}^2$, enquanto $f^{-1}(0)$ é a reunião dos dois eixos menos a origem, conjunto que não é fechado em $\mathbb{R}^2$ e tem quatro componentes conexas.

Exemplo 31. Uma matriz quadrada X, $n \times n$, chama-se *ortogonal* quando $X \cdot X^* = I$. Como se sabe da Álgebra Linear esta igualdade equivale a $X^* \cdot X = I$. Assim, $X \in \mathbb{R}^{n^2}$ é ortogonal se, e somente se, X^* é o inverso de X. Pela definição do produto de matrizes, $X \cdot X^* = I$ significa que os vetores-linha de X formam uma base ortonormal de $\mathbb{R}^n$, enquanto que $X^* \cdot X = I$ diz o mesmo sobre os vetores-coluna. Como estes são as imagens $X \cdot e_j$ dos vetores da base canônica de $\mathbb{R}^n$ pela transformação linear $X\colon \mathbb{R}^n \to \mathbb{R}^n$, vemos que X é ortogonal se, e somente se, transforma a base canônica de $\mathbb{R}^n$ (e portanto qualquer outra base ortonormal) numa base ortonormal. O conjunto $O(n) \subset \mathbb{R}^{n^2}$ das

matrizes ortogonais $n \times n$ constitui um grupo em relação à multiplicação usual de matrizes, chamado o *grupo ortogonal de* $\mathbb{R}^n$. Mostraremos agora que $O(n)$ é uma superfície C^∞ de dimensão $n(n-1)/2$ em $\mathbb{R}^{n^2}$. Para isto, começamos observando que, para todo $X \in \mathbb{R}^{n^2}$, o produto $X \cdot X^*$ é uma matriz simétrica e que o conjunto das matrizes simétricas é um espaço vetorial, naturalmente identificável a $\mathbb{R}^{n(n+1)/2}$, bastando que se desprezem, de cada matriz simétrica $X = (x_{ij})$, os elementos situados abaixo da diagonal principal, tomando como coordenadas de X os x_{ij} com $i \leq j$, isto é, os elementos acima da, ou sobre a, diagonal principal, os quais são em número de $1 + 2 + \cdots + n = n(n+1)/2$.

Seja, então, $f \colon \mathbb{R}^{n^2} \to \mathbb{R}^{n(n+1)/2}$ a aplicação de classe C^∞, definida por $f(X) = X \cdot X^*$. Para toda $X \in \mathbb{R}^{n^2}$, a derivada $f'(X)$ é a transformação linear $f'(X) \colon \mathbb{R}^{n^2} \to \mathbb{R}^{n(n+1)/2}$, dada por $f'(X) \cdot V = X \cdot V^* + V \cdot X^*$. Temos $O(n) = f^{-1}(I)$. Mostremos que I é um valor regular de f. Isto equivale a dizer que, para toda matriz ortogonal X e toda matriz simétrica S (ambas $n \times n$), existe uma matriz V, de ordem n, tal que $X \cdot V^* + V \cdot X^* = S$. Isto realmente ocorre: basta tomar $V = SX/2$. [Como adivinhar V? Pura sorte. Para ter a igualdade acima, a primeira tentativa é supor que cada parcela é igual a $S/2$. Pondo $V \cdot X^* = S/2$ e multiplicando por X obtemos logo $V = SX/2$.] Pelo Teorema 11, resulta que $O(n)$ é uma superfície de classe C^∞ e dimensão $n^2 - n(n+1)/2 = n(n-1)/2$ em $\mathbb{R}^{n^2}$. Observemos que $O(n)$ é compacta (já sabíamos isto desde o Capítulo I). O espaço vetorial tangente a $O(n)$ no ponto I, sendo o núcleo de $f'(I)$, é formado pelas matrizes V tais que $V + V^* = 0$, ou seja, pelas *matrizes anti-simétricas*, as quais, realmente, constituem um espaço vetorial de dimensão $n(n-1)/2$.

Sabemos que toda hiperfície é uma superfície de codimensão 1. A recíproca é verdadeira e decorre do

Teorema 12. *Toda superfície de classe C^k é localmente o gráfico de uma aplicação de classe C^k.*

Demonstração: Seja $M \subset \mathbb{R}^n$ uma superfície de classe C^k e dimensão m. Mostraremos que, para todo $p \in M$, existe uma parametrização $\varphi \colon V_0 \to V$, de classe C^k, definida num aberto $V_0 \subset \mathbb{R}^m$, com $p = \varphi(a) \in V$, a qual tem a forma $\varphi(x) = (x, f(x))$, onde $f \colon V_0 \to \mathbb{R}^{n-m}$ é uma aplicação de classe C^k. Assim, V será o gráfico de f. Começamos tomando uma parametrização arbitrária $\psi \colon W_0 \to W \subset M$, com $p = \psi(b) \in W$. Temos $T_pM = \psi'(b) \cdot \mathbb{R}^m$. Pelo Lema que antecede a demonstração do Teorema do Posto, existe uma decomposição em soma direta

$\mathbb{R}^n = \mathbb{R}^m \oplus \mathbb{R}^{n-m}$ tal que a correspondente projeção $\pi\colon \mathbb{R}^n \to \mathbb{R}^m$, na primeira coordenada, transforma T_pM isomorficamente sobre $\mathbb{R}^m$. Segue-se que a derivada $\pi \circ \psi'(b) = (\pi \circ \psi)'(b)\colon \mathbb{R}^m \to \mathbb{R}^m$ da aplicação $\pi \circ \psi\colon W_0 \to \mathbb{R}^m$, de classe C^k, é um isomorfismo. Pelo Teorema da Aplicação Inversa, existe um aberto $W_1 \subset \mathbb{R}^m$, com $b \in W_1 \subset W_0$, tal que $(\pi \circ \psi)|W_1$ é um difeomorfismo C^k sobre um aberto $V_0 \subset \mathbb{R}^m$. Seja $V = \psi(W_1)$. A aplicação $\varphi = \psi \circ [(\pi \circ \psi)|W_1]^{-1}\colon V_0 \to V$ é uma parametrização C^k. Além disso, para todo $x \in V_0$, $\pi\varphi(x) = x$, logo $\varphi(x) = (x, f(x))$, onde $f\colon V_0 \to \mathbb{R}^{n-m}$ é de classe C^k.

Corolário 1. *$M \subset \mathbb{R}^{n+1}$ é uma hiperfície de classe C^k se, e somente se, é uma superfície de codimensão 1 e classe C^k.*

Corolário 2. *Toda superfície $M \subset \mathbb{R}^n$, de classe C^k, é localmente a imagem inversa de um valor regular por uma aplicação de classe C^k.*

Com efeito, se $V \subset M$ é o gráfico de uma aplicação $f\colon V_0 \to \mathbb{R}^{n-m}$, de classe C^k no aberto $V_0 \subset \mathbb{R}^m$, então $V = g^{-1}(0)$, onde $g\colon V \times \mathbb{R}^{n-m} \to \mathbb{R}^{n-m}$ é definida por $g(x, y) = y - f(x)$.

Consideremos as três classes de subconjuntos do espaço euclidiano $\mathbb{R}^n$, abaixo descritas:

Classe A: formada pelos gráficos de aplicações de classe C^k, $f\colon V_0 \to \mathbb{R}^{n-m}$, definidas em abertos $V_0 \subset \mathbb{R}^m$;

Classe B: formada pelas imagens inversas de valores regulares de aplicações $g\colon Z \to \mathbb{R}^{n-m}$, de classe C^k em abertos $Z \subset \mathbb{R}^n$;

Classe C: formada pelas superfícies de classe C^k e dimensão m em $\mathbb{R}^n$.

Sabemos que $A \subset B \subset C$, que todo elemento de C está localmente em B e que cada elemento de B está localmente em A. Globalmente, porém, as três classes são distintas umas das outras. Em primeiro lugar, é evidente que $B \not\subset A$, pois A não contém conjuntos compactos (já que o gráfico de uma função contínua é sempre homeomorfo ao seu domínio e nenhum aberto não vazio $V_0 \subset \mathbb{R}^m$ é compacto). Também é verdade que $C \not\subset B$ mas não é tão evidente assim que haja alguma superfície que não pode ser definida como imagem inversa de um valor regular. Um exemplo é dado pelas superfícies não-orientáveis. Para explicar este ponto e, principalmente, porque se trata de uma noção indispensável para o estudo das integrais de superfícies, abordaremos a seguir o conceito de orientabilidade.

14 Superfícies orientáveis

Duas parametrizações $\varphi\colon V_0 \to V$, $\psi\colon W_0 \to W$, de classe C^1, numa superfície M, dizem-se *coerentes* quando $V \cap W = \varnothing$ ou então quando $V \cap W \neq \varnothing$ e a matriz jacobiana $J(\psi^{-1} \circ \varphi)(x)$ tem determinante positivo em todos os pontos $x \in \varphi^{-1}(V \cap W)$. Como sabemos, esta condição significa que $p = \varphi(x) = \psi(y) \in V \cap W$, e tem-se, para cada $j = 1, \ldots, m$,

$$\frac{\partial \varphi}{\partial x_j}(x) = \sum_{i=1}^{m} \alpha_{ij} \frac{\partial \psi}{\partial y_i}(y), \text{ com } \det(\alpha_{ij}) > 0.$$

Com efeito, (α_{ij}) é a matriz jacobiana do difeomorfismo $\psi^{-1} \circ \varphi$ no ponto x.

Se $V \cap W \neq \varnothing$, o difeomorfismo $\psi^{-1} \circ \varphi\colon \varphi^{-1}(V \cap W) \to \psi^{-1}(V \cap W)$ tem determinante jacobiano diferente de zero em todos os pontos $x \in \varphi^{-1}(V \cap W)$. Como $\det J(\psi^{-1} \circ \varphi)(x)$ é uma função contínua de x, seu sinal é constante em cada componente conexa do aberto $\varphi^{-1}(V \cap W) \subset \mathbb{R}^m$. Quando esse aberto é conexo, o sinal do determinante jacobiano é o mesmo em todos os pontos x. Convém entretanto observar que, em geral, $\varphi^{-1}(V \cap W)$ pode não ser conexo.

Um *atlas* de classe C^k numa superfície M é uma coleção $\mathcal{A}$ de parametrizações $\varphi\colon V_0 \to V$, de classe C^k cujas imagens cobrem M. O atlas $\mathcal{A}$ diz-se *coerente* quando quaisquer duas parametrizações $\varphi, \psi \in \mathcal{A}$ são coerentes. Um atlas coerente de classe C^k numa superfície M chama-se *máximo* quando não pode estar contido propriamente em nenhum outro atlas coerente de classe C^k em M. Todo atlas coerente $\mathcal{A}$ de classe C^k está contido num único atlas coerente máximo de classe C^k: basta acrescentar-lhe as parametrizações $\varphi\colon V_0 \to V$ de classe C^k que são coerentes com todas as $\psi \in \mathcal{A}$.

Uma superfície M de classe C^k diz-se *orientável* quando existe nela pelo menos um atlas coerente de classe C^k. Neste caso, existe também um atlas coerente máximo, o qual se chama uma *orientação* de M. Uma *superfície orientada* é uma superfície orientável na qual se fez a escolha de uma orientação $\mathcal{A}$. As parametrizações $\varphi \in \mathcal{A}$ chamam-se então *positivas*. Escreve-se $\varphi > 0$ para significar que φ pertence a $\mathcal{A}$.

Exemplo 32. Toda superfície que é a imagem de uma única parametrização é orientável, pois essa parametrização constitui um atlas coerente.

Isto inclui os subconjuntos abertos do espaço euclidiano e os gráficos de aplicações C^k definidas nesses abertos.

Exemplo 33. Todo subconjunto aberto U de uma superfície orientável M é uma superfície orientável: se $\mathcal{A}$ é um atlas coerente, as restrições $\varphi|\varphi^{-1}(U \cap V)$ das parametrizações $\varphi\colon V_0 \to V$, pertencentes a $\mathcal{A}$, constituem um atlas coerente em U.

Um *difeomorfismo* de classe C^k de uma superfície M sobre uma superfície N (ambas de classe C^k) é uma bijeção $f\colon M \to N$, de classe C^k, cuja inversa $f^{-1}\colon N \to M$ é diferenciável (e portanto de classe C^k). Pela Regra da Cadeia, se $f\colon M \to N$ é um difeomorfismo então, para todo $p \in M$, a derivada $f'(p)\colon T_pM \to T_qN$, $\; q = f(p)$, é um isomorfismo linear. [A recíproca é falsa: se cada $f'(p)$ é um isomorfismo, o Teorema da Aplicação Inversa garante apenas que f seja um difeomorfismo local.]

Se as superfícies M e N são difeomorfas então uma delas é orientável se, e somente se, a outra é. Com efeito, se $f\colon M \to N$ é um difeomorfismo de classe C^k então, para toda parametrização $\varphi\colon V_0 \to V$, de classe C^k, em N, a composta $f \circ \varphi\colon V_0 \to f(V)$ é uma parametrização C^k. As parametrizações do tipo $f \circ \varphi$ constituem um atlas em N e, como $(f \circ \psi)^{-1} \circ (f \circ \varphi) = \psi^{-1} \circ \varphi$, vemos que se φ percorre um atlas coerente em M, as compostas $f \circ \varphi$ constituem um atlas coerente em N. Portanto, M orientável $\Rightarrow$ N orientável. Usando f^{-1} em vez de f obteremos a outra implicação.

Uma observação simples, porém crucial, para a obtenção de um atlas coerente é a seguinte: dada uma parametrização $\varphi\colon V_0 \to V$, de classe C^k, numa superfície M, existe uma outra parametrização $\varphi_1\colon V_1 \to V$, também de classe C^k, *com a mesma imagem* V, tal que, para toda $\psi\colon W_0 \to W$, $\; \psi \in C^k$, em M, com $V \cap W \neq \varnothing$, tem-se $\det J(\psi^{-1} \circ \varphi_1)(x) = -\det J(\psi^{-1} \circ \varphi)(\bar{x})$, qualquer que seja $p = \varphi_1(x) = \varphi(\bar{x}) \in V \cap W$. Basta tomar $\varphi_1 = \varphi \circ \xi$, onde $\xi\colon V_1 \to V_0$ é qualquer difeomorfismo de classe C^k com determinante jacobiano negativo em todos os pontos. Por exemplo, podemos considerar a reflexão $\xi\colon \mathbb{R}^m \to \mathbb{R}^m$, dada por $\xi(x_1, \ldots, x_m) = (-x_1, x_2, \ldots, x_m)$ e pôr $V_1 = \xi^{-1}(V_0)$. Ao passar de φ para φ_1, diremos que estamos *mudando o sinal* da parametrização φ.

Como primeira aplicação da possibilidade de "mudar o sinal" de uma parametrização sem alterar sua classe de diferenciabilidade nem sua imagem temos o

Exemplo 34. Se uma superfície M admite um atlas formado por duas parametrizações $\varphi\colon V_0 \to V$, $\; \psi\colon W_0 \to W$, com $V \cap W$ conexo, então M

é orientável. Com efeito, como $\varphi^{-1}(V \cap W)$ é conexo, ou o determinante jacobiano de $\psi^{-1} \circ \varphi$ é positivo em todos os pontos, e então $\mathcal{A} = \{\varphi, \psi\}$ é um atlas coerente, ou é negativo em todos os pontos. Neste caso, mudando o sinal de φ, obtemos uma parametrização $\varphi_1 \colon V_1 \to V$ tal que o atlas $\{\varphi_1, \psi\}$ é coerente. Em qualquer caso, M é orientável. Usaremos agora este fato para mostrar que a esfera S^n é uma superfície orientável. Comecemos supondo $n \geq 2$. Sejam $p = e_{n+1}$ e $q = -p$ respectivamente o pólo norte e o pólo sul de S^n. Ponhamos $V = S^n - \{p\}$ e $W = S^n - \{q\}$. Temos $V \cup W = S^n$. No parágrafo 8 do Capítulo I, por meio de fórmulas explícitas, definimos a projeção estereográfica $\varphi \colon V \to \mathbb{R}^n$ e seu homeomorfismo inverso $\xi \colon \mathbb{R}^n \to V$. Vê-se, a partir daquelas fórmulas, que φ e ξ são aplicações de classe C^∞. Da igualdade $\varphi(\xi(y)) = y$, válida para todo $y \in \mathbb{R}^n$, resulta, via Regra da Cadeia, que $\xi'(y) \colon \mathbb{R}^n \to \mathbb{R}^{n+1}$ é injetiva em todos os pontos, logo ξ é uma parametrização C^∞. Definamos a parametrização $\zeta \colon \mathbb{R}^n \to W$ pondo $\zeta(y) = -\xi(y)$. Como $n \geq 2$, $V \cap W = S^n - \{p, q\}$ é conexo, portanto S^n é orientável. Quando $n = 1$, vemos que $S^1 - \{p, q\}$ não é conexo. Mas as parametrizações $\xi \colon \mathbb{R} \to S^1 - \{p\}$ e $\zeta \colon \mathbb{R} \to S^1 - \{q\}$, sendo dadas explicitamente por

$$\xi(t) = \left(\frac{2t}{t^2+1}, \frac{t^2-1}{t^2+1}\right) \quad \text{e} \quad \zeta(t) = \left(\frac{-2t}{t^2+1}, \frac{1-t^2}{t^2+1}\right),$$

observamos que $\zeta(-1/t) = \xi(t)$, logo a mudança de parametrização $\zeta^{-1} \circ \xi \colon \mathbb{R} - \{0\} \to \mathbb{R} - \{0\}$ tem a expressão $(\zeta^{-1} \circ \xi)(t) = -1/t$. O determinante jacobiano de $\zeta^{-1} \circ \xi$ é, neste caso, sua derivada, igual a $1/t^2$, portanto positiva para todo $t \in \mathbb{R} - \{0\}$. Assim, o atlas $\{\xi, \zeta\}$ é coerente e S^1 é orientável.

Seja $\mathcal{A}$ uma orientação numa superfície M de classe C^k. Como $\mathcal{A}$ é coerente, se $\varphi \colon V_0 \to V$ e $\psi \colon W_0 \to W$ são parametrizações positivas, tem-se $\det J(\psi^{-1} \circ \varphi)(x) > 0$ para todo $x \in \varphi^{-1}(V \cap W)$. E como $\mathcal{A}$ é máximo, se $\xi \colon V_0 \to V$ é uma parametrização de classe C^k não pertencente a $\mathcal{A}$, deve existir alguma $\psi \colon W_0 \to W$, com $\psi \in \mathcal{A}$, $V \cap W \neq \varnothing$ e algum ponto $p = \xi(x) \in V \cap W$ tal que $\det J(\psi^{-1} \circ \xi)(x) < 0$.

Teorema 13. *Sejam $\mathcal{A}$ um atlas coerente máximo de classe C^k na superfície M e $\xi \colon V_0 \to V$ uma parametrização de classe C^k em $V \subset M$, não pertencente ao atlas $\mathcal{A}$. Se V é conexo então, para toda parametrização $\psi \colon W_0 \to W$, pertencente a $\mathcal{A}$, tem-se $\det J(\psi^{-1} \circ \xi)(x) < 0$, seja qual for $x \in \xi^{-1}(V \cap W)$.*

Demonstração: O conjunto A, formado pelos pontos $p = \xi(x) \in V$ tais que existe uma parametrização $\psi\colon W_0 \to W$, pertencente a $\mathcal{A}$, com $p \in W$ e $\det J(\psi^{-1} \circ \xi)(x) < 0$ é aberto em V, como também é o conjunto B dos pontos $p = \xi(x)$ tais que existe $\varphi\colon Z_0 \to Z$, $\varphi \in \mathcal{A}$, $p \in Z$ e $\det J(\varphi^{-1} \circ \xi)(x) > 0$. Como $\mathcal{A}$ é um atlas, todo ponto de V está na imagem de alguma parametrização pertencente a $\mathcal{A}$, logo $V = A \cup B$. Além disso, se existisse $p = \xi(x) = \psi(y) \in A \cap B$, da igualdade $\varphi^{-1} \circ \xi = (\varphi^{-1} \circ \psi) \circ (\psi^{-1} \circ \xi)$ obteríamos

$$\det J(\varphi^{-1} \circ \xi)(x) = \det J(\varphi^{-1} \circ \psi)(y) \cdot \det J(\psi^{-1} \circ \xi)(x),$$

uma igualdade absurda pois o primeiro membro é positivo e, no segundo membro, o primeiro fator é positivo porque $\mathcal{A}$ é coerente e o segundo fator é negativo pela definição do conjunto A. Portanto $V = A \cup B$ é uma cisão. Mas V é conexo e $A \neq \varnothing$. (É neste ponto que usamos o fato de $\mathcal{A}$ ser máximo: de $\xi \notin \mathcal{A}$ segue-se que $A \neq \varnothing$ conforme a observação que precede o Exemplo 34.) Logo $B = \varnothing$, o que prova o teorema.

Corolário 1. *Se uma superfície M, de classe C^k, admite um atlas coerente de classe C^1 então admite também um atlas coerente de classe C^k.*

Com efeito, dado $\mathcal{A}$, um atlas coerente máximo de classe C^1 em M, seja $\mathcal{A}_k$ o conjunto das parametrizações $\varphi\colon V_0 \to V$ que são de classe C^k, pertencem ao atlas $\mathcal{A}$ e têm imagem V conexa. Como $\mathcal{A}_k \subset \mathcal{A}$, duas parametrizações quaisquer em $\mathcal{A}_k$ são coerentes. Além disso, como a superfície M é de classe C^k, dado qualquer ponto $p \in M$ existe uma parametrização $\xi\colon V_0 \to V$, de classe C^k, com V conexa e $p \in V$. Se $\xi \in \mathcal{A}$ então temos $\xi \in \mathcal{A}_k$. Se, porém, $\xi \notin \mathcal{A}$, pelo teorema acima teremos $\det J(\psi^{-1} \circ \xi)(x) < 0$ para toda $\psi\colon W_0 \to W$ em $\mathcal{A}$, com $V \cap W \neq \varnothing$, e todo $x \in \xi^{-1}(V \cap W)$. Neste caso, mudando o sinal de ξ obteremos uma parametrização $\xi_1\colon V_1 \to V$, de classe C^k, coerente com todas as parametrizações pertencentes a $\mathcal{A}$, logo $\xi_1 \in \mathcal{A}$ pela maximalidade deste atlas. Então $\xi_1 \in \mathcal{A}_k$, o que prova ser $\mathcal{A}_k$ um atlas.

O corolário acima esclarece e simplifica a exposição, mostrando que a classe de diferenciabilidade não influi no fato de ser a superfície M orientável. De agora por diante, salvo aviso em contrário, consideraremos apenas atlas coerentes de classe C^1. Quando for necessário, poderemos lançar mão de um atlas coerente da mesma classe que a superfície, autorizados pelo Corolário 1.

Dado um atlas coerente máximo $\mathcal{A}$ na superfície orientável M, diremos que uma parametrização $\xi\colon V_0 \to V$ em M é *negativa*, e escreveremos $\xi < 0$, quando, para toda parametrização $\psi\colon W_0 \to W$ pertencente ao atlas $\mathcal{A}$, e para todo ponto $p = \xi(x) \in V \cap W$, tivermos $\det J(\psi^{-1} \circ \xi)(x) < 0$. O teorema acima diz que se V é conexo então $\xi < 0$ se, e somente se, $\xi \notin \mathcal{A}$.

Duas parametrizações negativas são sempre coerentes: se $\xi\colon V_0 \to V$ e $\zeta\colon W_0 \to W$ são ambas negativas, para todo ponto $p \in V \cap W$ tome uma parametrização positiva φ cuja imagem contenha p e observe que $J(\zeta^{-1} \circ \xi) = J(\zeta^{-1} \circ \varphi) \cdot J(\varphi^{-1} \circ \xi)$. Assim, as parametrizações negativas constituem um atlas coerente $\mathcal{A}^*$, o qual é máximo, como se vê facilmente. A orientação de M definida pelo atlas $\mathcal{A}^*$ chama-se a *orientação oposta de* $\mathcal{A}$.

Corolário 2. *Sejam $\varphi\colon V_0 \to V$, $\psi\colon W_0 \to W$ parametrizações de classe C^1 uma superfície M, com V, W conexas e $V \cap W \neq \varnothing$. Se o determinane jacobiano de $\psi^{-1} \circ \varphi\colon \varphi^{-1}(V \cap W) \to \psi^{-1}(V \cap W)$ muda de sinal então M não é orientável.*

Com efeito, se M fosse orientável, tomaríamos um atlas coerente máximo $\mathcal{A}$, de classe C^1 em M. A hipótese feita impede $\varphi, \psi \in \mathcal{A}$. O teorema acima também impede que uma dessas parametrizações pertença a $\mathcal{A}$ e a outra não. Restaria apenas a possibilidade de $\varphi \notin \mathcal{A}$ e $\psi \notin \mathcal{A}$. Como V e W são conexas, teríamos $\varphi < 0$ e $\psi < 0$, logo φ e ψ seriam coerentes, como observamos acima, mas isto contraria a hipótese do corolário. Logo M é não-orientável.

Observação: O determinante jacobiano de $\psi^{-1} \circ \varphi$ nunca se anula. Logo ele só pode mudar de sinal quando a interseção $V \cap W$ é desconexa.

Exemplo 35. Seja $M \subset \mathbb{R}^6$ o conjunto das matrizes 2×3 de posto 1. Escrevendo cada matriz 2×3 sob a forma (v, w), onde $v, w \in \mathbb{R}^3$ são suas linhas, M é o conjunto das matrizes $(v, w) \in \mathbb{R}^6 - \{0\}$ tais que v e w são linearmente dependentes. Temos $M = V \cup W$, onde V é o conjunto das matrizes de M cuja primeira linha é $\neq 0$ enquanto as matrizes de W têm a segunda linha não nula. Sejam $U = \mathbb{R}^3 - \{0\}$ e $V_0 = U \times \mathbb{R}$. As aplicações $\varphi\colon V_0 \to V$, $\psi\colon V_0 \to W$, definidas por $\varphi(v, t) = (v, tv)$ e $\psi(v, t) = (tv, v)$, são parametrizações de classe C^∞ e $\varphi^{-1}(V \cap W) = U \times (\mathbb{R} - \{0\}) = \psi^{-1}(V \cap W)$ tem duas componentes conexas, a saber, $U \times (-\infty, 0)$ e $U \times (0, +\infty)$. A mudança de parametrização $\psi^{-1} \circ \varphi\colon (v, t) \mapsto (tv, 1/t)$, ou seja, $(x, y, z, t) \mapsto (tx, ty, tz, 1/t)$, tem

determinante jacobiano igual a $-t$, o qual é positivo na primeira componente conexa e negativo na segunda. M é portanto uma superfície de classe C^∞ e dimensão 4 em $\mathbb{R}^6$, não-orientável pelo Corolário 2 acima.

Exemplo 36. Se M e N são superfícies orientáveis, o produto cartesiano $M \times N$ é também orientável. Com efeito, se $\mathcal{A}$ e $\mathcal{B}$ são atlas coerentes em M e N respectivamente, seja $\mathcal{A} \times B$ o atlas em $M \times N$ formado pelas parametrizações $\varphi \times \psi$, onde $\varphi\colon V_0 \to V$ pertence a $\mathcal{A}$, $\psi\colon W_0 \to W$ pertence a $\mathcal{B}$ e $\varphi \times \psi\colon V_0 \times W_0 \to V \times W$ é definida por $(\varphi \times \psi)(x, y) = (\varphi(x), \psi(y))$. Como se vê facilmente, $(\varphi_2 \times \psi_2)^{-1} \circ (\varphi_1 \times \psi_1) = (\varphi_2^{-1} \circ \varphi_1) \times (\psi_2^{-1} \circ \psi_1)$, donde

$$\det J[(\varphi_2 \times \psi_2)^{-1} \circ (\varphi_1 \times \psi_1)](x, y) =$$
$$= \det J(\varphi_2^{-1} \circ \varphi_1)(x) \cdot \det J(\psi_2^{-1} \circ \psi_1)(y).$$

Isto mostra que $\mathcal{A} \times \mathcal{B}$ é um atlas coerente em $M \times N$.

Reciprocamente, se $M \times N$ é orientável, então M e N são orientáveis. Com efeito, seja $\mathcal{C}$ um atlas coerente máximo em $M \times N$ e consideremos a coleção $\mathcal{A}$ das parametrizações φ em M tais que $\varphi \times \zeta \in \mathcal{C}$, onde $\zeta\colon Z_0 \to Z$ é uma parametrização em N, com imagem Z conexa, a qual fixamos de uma vez por todas. Se $\varphi\colon V_0 \to V$ e $\psi\colon W_0 \to W$ pertencem a $\mathcal{A}$ e $V \cap W \neq \varnothing$, pondo $\alpha = (\psi \times \zeta)^{-1} \circ (\varphi^{-1} \times \zeta)$, vemos que $\alpha = (\psi^{-1} \circ \varphi) \times \mathrm{id} : \varphi^{-1}(V \cap W) \times Z_0 \to \psi^{-1}(V \cap W) \times Z_0$, logo $\det J\alpha(x, y) = \det J(\psi^{-1} \circ \varphi)(x)$. Como o atlas $\mathcal{C}$ é coerente, temos $\det J\alpha(x, y) > 0$ para todo (x, y). Assim $\det J(\psi^{-1} \circ \varphi)(x) > 0$ para todo $x \in \varphi^{-1}(V \cap W)$, ou seja, duas parametrizações em $\mathcal{A}$ são coerentes. Além disso, dado qualquer $p \in M$, tomamos uma parametrização $\varphi\colon V_0 \to V$ em M, com $p = \varphi(x) \in V$ e V conexo. Se $\varphi \notin \mathcal{A}$ então $\varphi \times \zeta < 0$. Mudando o sinal de φ, obtemos uma parametrização $\varphi_1 : V_1 \to V$, com a mesma imagem que φ, tal que $\varphi_1 \times \zeta \in \mathcal{C}$. Portanto $\mathcal{A}$ é um atlas e M é orientável. De modo análogo se mostra que N também é orientável.

Resulta da última parte do Exemplo 36 uma outra maneira de verificar que, para todo $n \geq 1$, a esfera S^n é orientável. Com efeito, a aplicação $h\colon S^n \times \mathbb{R} \to \mathbb{R}^{n+1} - \{0\}$, definida por $h(x, t) = e^t \cdot x$, é um difeomorfismo do produto cartesiano $S^n \times \mathbb{R}$ sobre um subconjunto aberto do espaço euclidiano $\mathbb{R}^{n+1}$. [O difeomorfismo inverso, $h^{-1}\colon \mathbb{R}^{n+1} - \{0\} \to S^n \times \mathbb{R}$ é dado por $h^{-1}(z) = (z/|z|, \log|z|)$.] Portanto o produto cartesiano $S^n \times \mathbb{R}$ é orientável e daí decorre que S^n é orientável.

Talvez a maneira mais popular de se caracterizar uma superfície orientável seja por meio de campos de vetores normais. Para superfícies de codimensão $n-m \geq 1$, a existência de $n-m$ campos contínuos de vetores normais linearmente independentes é uma condição suficiente da orientabilidade, sendo também necessária em codimensão 1.

Dada uma superfície $M \subset \mathbb{R}^n$, o *espaço vetorial normal* a M no ponto p é o conjunto $T_pM^\perp$ dos vetores $w \in \mathbb{R}^n$ tais que $\langle w, v\rangle = 0$ para todo $v \in T_pM$, ou seja, é o complemento ortogonal do espaço vetorial tangente a M no ponto p. Se $\dim M = m$ então $\dim T_pM^\perp = n - m = \operatorname{codim}.M$. Os elementos $w \in T_pM^\perp$ são chamados os *vetores normais* a M no ponto p.

Por exemplo, o espaço vetorial normal à esfera S^n no ponto p é o conjunto dos vetores $w = t \cdot p,\ \ t \in \mathbb{R}$. Mais geralmente, se $c \in \mathbb{R}^{n-m}$ é valor regular da aplicação $f\colon U \to \mathbb{R}^{n-m}$, de classe C^1 no aberto $U \subset \mathbb{R}^n$, então a superfície $M = f^{-1}(c)$ tem em cada um dos seus pontos p o espaço vetorial normal gerado pelos $n-m$ vetores linearmente independentes $\operatorname{grad} f_1(p), \ldots, \operatorname{grad} f_{n-m}(p)$, onde $f_1, \ldots, f_{n-m}\colon U \to \mathbb{R}$ são as funções-coordenada de f. No caso da esfera, temos $f\colon \mathbb{R}^{n+1} \to \mathbb{R}$ definida por $f(x) = \langle x, x\rangle,\ \ S^n = f^{-1}(1)$ e $\operatorname{grad} f(x) = 2x$. Logo o vetor $2p$ (e portanto o vetor p) é uma base do espaço vetorial normal a S^n no ponto p.

Um *campo de vetores normais* à superfície $M \subset \mathbb{R}^n$ é uma aplicação $v\colon M \to \mathbb{R}^n$ tal que, para todo $p \in M,\ \ v(p) \in T_pM^\perp$. Os adjetivos "contínuo", "diferenciável", "de classe C^k", etc., aplicados ao campo, são propriedades da aplicação v.

Teorema 14. *Se uma superfície $M \subset \mathbb{R}^n$, de dimensão m, admite $n-m$ campos contínuos de vetores normais $v_1, \ldots, v_{n-m} : M \to \mathbb{R}^n$ linearmente independentes em cada ponto $p \in M$, então M é orientável.*

Demonstração: Seja $\mathcal{A}$ o conjunto das parametrizações de classe C^1, $\varphi\colon V_0 \to V$, em M, tais que V_0 é conexo e, para todo $x \in V_0$, a matriz $n \times n$

$$\Phi(x) = \left[\frac{\partial\varphi}{\partial x_1}(x), \ldots, \frac{\partial\varphi}{\partial x_m}(x),\, v_1(\varphi(x)), \ldots, v_{n-m}(\varphi(x))\right],$$

cujas colunas são os vetores aí indicados, tem determinante positivo. Para cada $x \in V_0$, note que $\Phi(x)$ é invertível pois suas primeiras m colunas formam uma base de $T_{\varphi(x)}M$ e as restantes formam uma base do complemento ortogonal desse subespaço de $\mathbb{R}^n$. Como $\Phi(x)$ depende

continuamente de x, seu determinante não muda de sinal no conjunto conexo V_0. Se, para uma certa parametrização $\varphi\colon V_0 \to V$ (V_0 conexo), tivermos $\det \Phi(x) < 0$, podemos trocar o sinal de φ e obter $\varphi_1 \in \mathcal{A}$, com a mesma imagem V. Logo $\mathcal{A}$ é um atlas de classe C^1 em M. Para mostrar que $\mathcal{A}$ é coerente, sejam $\varphi\colon V_0 \to V$, $\psi\colon W_0 \to W$ pertencentes a $\mathcal{A}$ e $p = \varphi(x) = \psi(y) \in V \cap W$. Sabemos que $J(\psi^{-1} \circ \varphi)(x) = (\alpha_{ij}) = A$ onde, para cada $j = 1, \ldots, m$,

$$\frac{\partial \varphi}{\partial x_j}(x) = \sum_{i=1}^{m} \alpha_{ij} \frac{\partial \psi}{\partial y_i}(y),$$

Isto significa, em termos de matrizes, que $\Phi(x) = \Psi(y) \cdot \widetilde{A}$, com $\widetilde{A} = \begin{pmatrix} A & 0 \\ 0 & I \end{pmatrix}$, onde I é a matriz identidade de ordem $n-m$. Como $\det \Phi(x) > 0$ e $\det \Psi(y) > 0$ resulta que $0 < \det \widetilde{A} = \det A = \det J(\psi^{-1} \circ \varphi)(x)$. Isto conclui a demonstração.

Corolário. *Se a superfície $M \subset \mathbb{R}^n$ pode ser definida como imagem inversa de um valor regular de uma aplicação $f\colon U \to \mathbb{R}^{n-m}$, de classe C^1 num aberto $U \subset \mathbb{R}^n$, então M é orientável.*

Exemplo 37. Daí resulta, por exemplo, que a superfície de dimensão 4 em $\mathbb{R}^6$, formada pelas matrizes 2×3 de posto 1 (veja Exemplo 35) não pode ser definida como imagem inversa de um valor regular de uma aplicação $f\colon U \to \mathbb{R}^2$, de classe C^1 num aberto $U \subset \mathbb{R}^6$, pois não é orientável. Por outro lado, o grupo ortogonal $O(n)$ é uma superfície orientável de dimensão $n(n-1)/2$ em $\mathbb{R}^{n^2}$. (Veja Exemplo 31.)

Exemplo 38. A existência de r campos contínuos de vetores normais linearmente independentes em todos os pontos de uma superfície $M \subset \mathbb{R}^n$ de codimensão $n - m > r$ não garante que M seja orientável. Para ver isto, basta tomar $M \subset \mathbb{R}^n$ e passar a considerar $M \subset \mathbb{R}^n \times \mathbb{R}^r = \mathbb{R}^{n+r}$. (As r últimas coordenadas de cada ponto $p \in M$, considerado como ponto em $\mathbb{R}^{n+r}$, são nulas.) Então os r campos constantes $e_{n+1}, \ldots, e_{n+r}\colon M \to \mathbb{R}^{n+r}$ são contínuos, linearmente independentes e normais a M, que pode perfeitamente ter sido tomada não-orientável.

Quando a codimensão de M é igual a 1, vale a recíproca do teorema acima.

Teorema 15. *Uma hiperfície $M \subset \mathbb{R}^{n+1}$ é orientável se, e somente se, existe um campo contínuo de vetores normais $v\colon M \to \mathbb{R}^{n+1}$, com $v(p) \neq 0$ para todo $p \in M$.*

Demonstração: O campo de vetores normais não nulos $v\colon M \to \mathbb{R}^{n+1}$ em M determina, em cada ponto $p \in M$, uma base $\{v(p)\}$ do espaço vetorial normal $T_pM^\perp$, que é uma reta. Portanto a condição é suficiente, pelo Teorema 14. Reciprocamente, se M é orientável, definiremos um campo contínuo de vetores normais unitários $u\colon M \to \mathbb{R}^{n+1}$ do seguinte modo. Para cada $p \in M$, existem dois vetores unitários normais a M no ponto p. Escolhemos qual deles é $u(p)$ tomando uma parametrização positiva $\varphi\colon V_0 \to V$, com $p = \varphi(x) \in V$, e definindo $u(p)$ pela condição de que a matriz

$$\Phi(x) = \left[\frac{\partial\varphi}{\partial x_1}(x), \ldots, \frac{\partial\varphi}{\partial x_n}(x),\, u(\varphi(x))\right],$$

cujas $n+1$ colunas são os vetores aí indicados, tenha determinante positivo. Isto não depende da parametrização positiva tomada porque se $\psi\colon W_0 \to W$, com $p = \psi(y) \in W$ é outra delas, então $\Phi(x) = \Psi(y)\cdot\widetilde{A}$, onde $\widetilde{A} = \begin{pmatrix} A & 0 \\ 0 & I \end{pmatrix}$ e $A = (\alpha_{ij})$ é a matriz jacobiana de $\psi^{-1}\circ\varphi$ no ponto x, a qual, como sabemos, cumpre a condição

$$\frac{\partial\varphi}{\partial x_j}(x) = \sum_{i=1}^{m} \alpha_{ij}\,\frac{\partial\psi}{\partial y_i}(y).$$

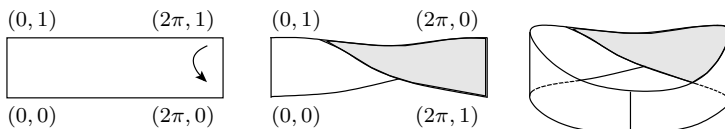
Como φ e ψ são coerentes, temos $\det A > 0$, ou seja, $\det \widetilde{A} > 0$. Assim $\det \Phi(x) > 0$ se, e somente se, $\det \Psi(y) > 0$. Resta mostrar que o campo u é contínuo. Para tal, observamos que (como toda superfície) M é localmente a imagem inversa $f^{-1}(c)$ de um valor regular. Logo M pode ser coberta por vizinhanças conexas parametrizadas V, em cada uma das quais está definido um campo contínuo de vetores unitários normais $v\colon V \to \mathbb{R}^{n+1}$, a saber $v(p) = \operatorname{grad} f(p)/|\operatorname{grad} f(p)|$. Dada uma parametrização positiva $\varphi\colon V_0 \to V$, ou se tem

$$\det\left[\frac{\partial\varphi}{\partial x_1}(x), \ldots, \frac{\partial\varphi}{\partial x_n}(x),\, v(\varphi(x))\right] > 0$$

para todo $x \in V_0$, ou então esse determinante é negativo em todos os pontos de V_0. No primeiro caso, temos $v(p) = u(p)$ para todo

$p = \varphi(x) \in V$ e, no segundo, $v(p) = -u(p)$ para todo $p = \varphi(x)$. Em qualquer caso, u é contínuo em V. Como as vizinhanças V cobrem M, concluimos que o campo $u\colon M \to \mathbb{R}^{n+1}$ é contínuo.

Observação: Podemos, sempre que seja necessário, supor que o campo v é unitário (isto é, $|v(p)| = 1$ para todo $p \in M$) substituindo-o, se for o caso, pelo campo w, onde $w(p) = v(p)/|v(p)|$.

Exemplo 39. A superfície não-orientável mais conhecida é sem dúvida, a *faixa de Moebius*. Ela pode ser obtida do retângulo $[0, 2\pi] \times (0, 1)$ colando-se, para cada $t \in (0, 1)$ os pontos $(0, t)$ e $(2\pi, 1 - t)$.

Podemos também definir a faixa de Moebius como a superfície $M \subset \mathbb{R}^3$, obtida pela rotação de um segmento de reta aberto, cujo ponto médio se apoia sobre um círculo. Enquanto o ponto médio desliza sobre o círculo, o segmento realiza uma rotação de $180°$ até o final da primeira volta. Para parametrizar M, consideremos a aplicação $f\colon (0, 1) \times \mathbb{R} \to \mathbb{R}^3$ de classe C^∞, definida por $f(s, t) = \gamma(t) + (s - 1/2)\delta(t)$, onde $\gamma(t) = (\cos t, \operatorname{sen} t, 0)$ e $\delta(t) = \cos(t/2) \cdot \gamma(t) + \operatorname{sen}(t/2) \cdot e_3$. A imagem de f é a faixa de Moebius M. Para cada intervalo aberto $I \subset \mathbb{R}$, de comprimento $\leq \pi$, a restrição de f a $(0, 1) \times I$ é uma parametrização em M, logo M é uma superfície de classe C^∞ em $\mathbb{R}^3$. Para mostrar que M não é orientável, façamos primeiro uma observação mais geral: se $M \subset \mathbb{R}^{n+1}$ é uma hiperfície orientável e $\lambda\colon [a, b] \to M$ é um caminho fechado em M então, dando qualquer "campo contínuo $u\colon [a, b] \to \mathbb{R}^n$ de vetores unitários normais a M ao longo de λ", deve-se ter necessariamente $u(a) = u(b)$. [A expressão entre aspas significa que a aplicação u é contínua e, para cada $t \in [a, b]$, o vetor $u(t) \in \mathbb{R}^n$ tem comprimento 1 e é normal a M no ponto $\lambda(t)$.] Para provar esta afirmação consideremos o campo contínuo $v\colon M \to \mathbb{R}^n$, de vetores unitários normais a M, o qual existe em virtude do teorema acima. Para cada $t \in [a, b]$, temos $\langle v(\lambda(t)), u(t)\rangle = \pm 1$. Como este produto interno depende continuamente de t no conjunto conexo $[a, b]$, ele deve ser constante, igual a 1 ou a -1 para todo t. Em particular, $\langle v(\lambda(a)), u(a)\rangle = \langle v(\lambda(b)), u(b)\rangle$. Sendo $\lambda(a) = \lambda(b)$, isto obriga $u(a) = u(b)$, como foi afirmado. Ora,

ao longo do caminho fechado $\gamma\colon [0,2\pi] \to M$, cintura da faixa de Moebius, podemos definir o campo contínuo de vetores normais unitários $u\colon [0,2\pi] \to \mathbb{R}^3$, dado por

$$u(t) = \left(-\cos t \cdot \operatorname{sen} \frac{t}{2}\,,\; -\operatorname{sen} t \cdot \operatorname{sen} \frac{t}{2}\,,\; \cos \frac{t}{2}\right).$$

Temos $u(2\pi) = -u(0)$, logo M não é orientável. [Para chegar à fórmula explícita das coordenadas de $u(t)$, basta notar que

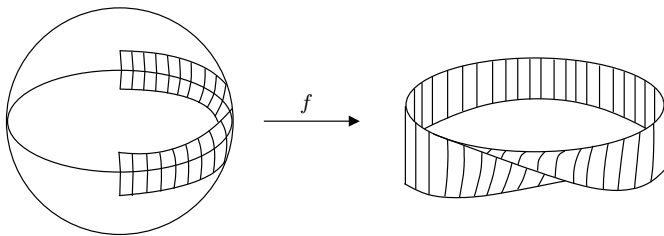
$$u(t) = \frac{\partial f}{\partial s}(1/2,t) \times \frac{\partial f}{\partial t}(1/2,t)$$

é o *produto vetorial* dos vetores tangentes básicos determinados pela parametrização f no ponto $f(1/2,t) = \gamma(t)$.]

Exemplo 40. *O plano projetivo como superfície de codimensão* 2 *em* $\mathbb{R}^4$. Seja $f\colon \mathbb{R}^3 \to \mathbb{R}^4$ a aplicação de classe C^∞ definida por $f(x,y,z) = (x^2-y^2, xy, xz, yz)$. Afirmamos que o conjunto $P^2 = f(S^2)$, imagem por f da esfera unitária $S^2 \subset \mathbb{R}^3$, é uma superfície de dimensão 2 e classe C^∞ em $\mathbb{R}^4$. Em primeiro lugar, o leitor pode verificar sem maiores dificuldades que, se $p,q \in S^2$, então $f(p) = f(q)$ se, e somente se, $q = \pm p$. (Quando $q = -p$, diz-se que os pontos p e q, pertencentes a S^2, são *antípodas*.) Em seguida, mostraremos que, em cada ponto $p = (x,y,z) \in S^2$, a derivada $f'(p)\colon \mathbb{R}^3 \to \mathbb{R}^4$ aplica o plano tangente $T_pS^2 \subset \mathbb{R}^3$ injetivamente em $\mathbb{R}^4$. Com efeito, a matriz jacobiana de f no ponto p é

$$Jf(x,y,z) = \begin{pmatrix} 2x & -2y & 0 \\ y & x & 0 \\ z & 0 & x \\ 0 & z & y \end{pmatrix}.$$

Dois dos determinantes menores 3×3 de Jf são $2x(x^2+y^2)$ e $2y(x^2+y^2)$. Logo Jf tem posto 3 exceto quando $x = y = 0$. Portanto $f'(p)\colon \mathbb{R}^3 \to \mathbb{R}^4$ é injetiva para todo $p \in S^2 - \{a,-a\}$, $a = (0,0,1)$. Nos pontos $\pm a$, a matriz jacobiana mostra que $f'(\pm a)\cdot e_1 = \pm e_3$ e $f'(\pm a)\cdot e_2 = \pm e_4$. Como os planos tangentes a S^2 nos pontos $\pm a$ coincidem e têm por base $\{e_1, e_2\}$, vemos que $\dim[f'(\pm a)\cdot T_{\pm a}(S^2)] = 2$. Segue-se que $\dim[f'(p)\cdot T_p(S^2)] = 2$ para todo $p \in S^2$. Daí resulta que, para cada parametrização $\varphi\colon V_0 \to V$, de classe C^∞, de um aberto $V \subset S^2$, a aplicação $f\circ\varphi\colon V_0 \to \mathbb{R}^4$ é uma imersão C^∞. Se V não contiver um

par de pontos antípodas, $f \circ \varphi$ será injetiva. Afirmamos que, neste caso, $f \circ \varphi$ será um homeomorfismo de V_0 sobre um subconjunto $W \subset P^2$, *aberto* em P^2. Isto é verdade porque a restrição $f|S^2 \colon S^2 \to P^2$ é uma aplicação aberta, fato que passamos a demonstrar: dado um subconjunto aberto $A \subset S^2$, devemos mostrar que $f(A)$ é aberto em P^2, ou seja, que $P^2 - f(A)$ é fechado. Pela observação (12.5), §12 do Capítulo I, basta mostrar que $f^{-1}(P^2 - f(A)) = S^2 - f^{-1}f(A) = S^2 - [A \cup (-A)]$ é fechado, ou seja, que $A \cup (-A)$ é aberto em S^2, o que é óbvio pois A e $-A = \{x \in S^2; -x \in A\}$ são abertos em S^2. Portanto $P^2 = f(S^2)$ é uma superfície bidimensional de classe C^∞ em $\mathbb{R}^4$. Pelo fato de e-xistir uma aplicação contínua sobrejetiva $f \colon S^2 \to P^2$ tal que $f(p) = f(q) \Leftrightarrow q = \pm p$, P^2 diz-se um modelo do *plano projetivo* em $\mathbb{R}^4$. A superfície P^2 é compacta, por ser uma imagem contínua do compacto S^2. Por outro lado, P^2 não é orientável porque contém um conjunto aberto homeomorfo a uma faixa de Moebius, a saber, a imagem por f da região $R = \{(x, y, z) \in S^2; y \geq 0, -1/2 < z < 1/2\}$. (Observe que f é um homeomorfismo no interior de R relativamente a S^2 mas no bordo ∂R, formado por dois arcos de círculo, por causa da relação $f(p) = f(-p)$, f realiza uma colagem como a que dá origem à faixa de Moebius).

A noção de superfície orientável também pode ser apresentada a partir do conceito de orientação de um espaço vetorial. Em determinados contextos, essa abordagem é mais vantajosa. Vamos descrevê-la a seguir.

Seja E um espaço vetorial (real) de dimensão $m > 0$. Dadas duas bases (ordenadas!) $\mathcal{E} = \{e_1, \ldots, e_m\}$ e $\mathcal{F} = \{f_1, \ldots, f_m\}$ em E, existe uma única matriz real $m \times m$ invertível $A = (a_{ij})$ tal que

$$f_j = \sum_{i=1}^{m} a_{ij} e_i$$

para todo $j = 1, \ldots, m$. A chama-se a *matriz de passagem* da base $\mathcal{E}$ para a base $\mathcal{F}$. Diremos que as bases $\mathcal{E}$ e $\mathcal{F}$ são *igualmente orientadas*

quando a matriz de passagem de $\mathcal{E}$ para $\mathcal{F}$ tiver determinante positivo. Escreveremos então $\mathcal{E} \equiv \mathcal{F}$. Esta é uma relação de equivalência no conjunto das bases de E. Cada classe de equivalência segundo esta relação chama-se uma *orientação* no espaço vetorial E.

Seja $\mathcal{O}$ uma orientação no espaço vetorial E. $\mathcal{O}$ é um conjunto de bases (ordenadas) de E com as seguintes propriedades. Primeiro, duas bases quaisquer em $\mathcal{O}$ são igualmente orientadas. Segundo, se $\mathcal{E}$ pertence a $\mathcal{O}$ e $\mathcal{F} \equiv \mathcal{E}$ então $\mathcal{F}$ pertence a $\mathcal{O}$. Assim, a orientação $\mathcal{O}$ fica determinada por qualquer um dos seus elementos $\mathcal{E}$: $\mathcal{O}$ consiste de todas as bases de E igualmente orientadas em relação a $\mathcal{E}$. Se as matrizes de passagem de $\mathcal{F}$ para as bases $\mathcal{F}$ e $\mathcal{G}$ são respectivamente A e B então, como se vê facilmente, a matriz de passagem de $\mathcal{E}$ para $\mathcal{G}$ é BA^{-1}. Se det. $A < 0$ e det. $B < 0$ então det. $BA^{-1} > 0$. Ou seja, se $\mathcal{F}$ e $\mathcal{G}$ não pertencem à orientação $\mathcal{O}$ então $\mathcal{F} \equiv \mathcal{G}$. Isto mostra que a relação $\equiv$ possui duas classes de equivalência. Noutras palavras, o espaço vetorial E admite duas orientações.

Um *espaço vetorial orientado* é um par $(E, \mathcal{O})$, onde $\mathcal{O}$ é uma orientação em E. Muitas vezes escreveremos apenas E, deixando subentendida a orientação $\mathcal{O}$. Uma vez escolhida $\mathcal{O}$, a outra orientação de E será chamada a *orientação oposta* e a indicaremos com $-\mathcal{O}$. No espaço vetorial orientado E, as bases pertencentes a $\mathcal{O}$ serão chamadas *positivas* e as outras *negativas*.

Usando o fato de que o conjunto das matrizes $m \times m$ com determinante positivo é conexo por caminhos (veja o Apêndice ao Capítulo VII), podemos dar a seguinte interpretação topológica para a relação $\mathcal{E} \equiv \mathcal{F}$. Se temos

$$f_j = \sum_{i=1}^{m} a_{ij} e_i \ \ (j = 1, \ldots, m),$$

tomemos um caminho de matrizes $A(t) = (a_{ij}(t))$, $t \in [0, 1]$, com $A(0) =$ matriz identidade $m \times m$, $A(1) = (a_{ij})$, det $.A(t) > 0$ para todo t, e ponhamos $h_j(t) = \sum_{i=1}^{m} a_{ij}(t) e_i$ $(j = 1, \ldots, m)$. Então $\mathcal{H}(t) = \{h_1(t), \ldots, h_m(t)\}$ é, para cada $t \in [0, 1]$, uma base de E, positivamente orientada em relação a $\mathcal{E}$, com $\mathcal{H}(0) = \mathcal{E}$ e $\mathcal{H}(1) = \mathcal{F}$. Assim, se duas bases de E são igualmente orientadas, pode-se deformar uma delas continuamente até obter a outra, de modo que em cada instante t da deformação se tenha uma base de E. A recíproca é evidente. [As questões de continuidade no espaço vetorial m-dimensional E se referem

a uma norma qualquer que nele tomemos. Não importa qual norma, pois todas são equivalentes. Nos casos que consideraremos, E será sempre um subespaço de $\mathbb{R}^n$.]

Sejam E, F espaços vetoriais orientados, de mesma dimensão. Um isomorfismo $T\colon E \to F$ diz-se *positivo* quando transforma bases positivas de E em bases positivas de F. Para que isto ocorra, basta que T transforme *uma* base positiva de E numa base positiva de F. Diz-se também, neste caso, que T *preserva orientação.* No caso contrário, diz-se que T é *negativo*, ou que T *inverte orientação.* Evidentemente, se $T\colon E \to F$ preserva orientação então $T^{-1}\colon F \to E$ também preserva. Além disso, se $S\colon E \to F$ e $T\colon F \to G$ preservam orientação, o mesmo se dá com $T \cdot S\colon E \to G$.

O espaço $\mathbb{R}^m$ será considerado orientado pela exigência de que sua base canônica seja positiva. Uma transformação linear $T\colon \mathbb{R}^m \to \mathbb{R}^m$ é um isomorfismo positivo se, e somente se, $\det .T > 0$.

Se apenas um dos espaços vetoriais E, F é orientado, a exigência de que o isomorfismo $T\colon E \to F$ seja positivo determina univocamente uma orientação no outro espaço.

Seja M uma superfície orientada de dimensão m. Em cada ponto $x \in M$ definimos uma orientação $\mathcal{O}_x$ no espaço vetorial tangente T_xM declarando que se $\varphi\colon U_0 \to U$ é uma parametrização positiva em M, com $x = \varphi(u) \in U$ então a base $\left\{\frac{\partial\varphi}{\partial u_1}(u), \ldots, \frac{\partial\varphi}{\partial u_m}(u)\right\} \subset T_xM$ pertence a $\mathcal{O}_x$, o que equivale a dizer que o isomorfismo $\varphi'(u)\colon \mathbb{R}^m \to (T_xM, \mathcal{O}_x)$ preserva orientação. Resulta da definição de atlas coerente que a orientação $\mathcal{O}_x$ não depende da parametrização positiva φ escolhida para defini-la. Em geral, uma base $\{w_1, \ldots, w_m\} \subset T_xM$ é positiva se, e somente se, $w_j = \sum_{i=1}^{m} a_{ij} \frac{\partial\varphi}{\partial u_i}(u)$, $x = \varphi(u)$, onde $\varphi > 0$ e $\det(a_{ij}) > 0$.

Reciprocamente, dada a superfície M, suponhamos que foi escolhida, para cada $x \in M$, uma orientação $\mathcal{O}_x$ no espaço vetorial tangente T_xM. Suponhamos ainda que esta escolha foi feita de tal modo que todo ponto de M pertença à imagem de uma parametrização $\varphi\colon U_0 \to U$ cuja derivada $\varphi'(u)\colon \mathbb{R}^m \to (T_xM, \mathcal{O}_x)$, $x = \varphi(u)$, preserva orientação, para todo $u \in U_0$. Evidentemente, se $\psi\colon V_0 \to V$ é outra parametrização em M, tal que $\psi'(v)\colon \mathbb{R}^m \to (T_yM, \mathcal{O}_y)$ preserva orientação para todo $v \in V_0$, com $y = \psi(v)$, e $U \cap V \neq \varnothing$ então $\psi^{-1} \circ \varphi\colon \varphi^{-1}(U \cap V) \to \psi^{-1}(U \cap V)$ tem determinante jacobiano positivo em todos os pontos do seu domínio. Logo, M é uma superfície orientável.

Sejam M, N superfícies orientadas de mesma dimensão m e $f: M \to N$ um difeomorfismo local de classe C^1. (Isto significa que M pode ser coberta por abertos U, restrita a cada um dos quais f é um difeomorfismo sobre um aberto de N. Pelo Teorema da Aplicação Inversa, podemos dizer, equivalentemente, que a derivada $f'(x): T_xM \to T_{f(x)}N$ é, em cada ponto $x \in M$, um isomorfismo.)

Dizemos que f *preserva orientação* (ou é *positivo*) quando, para cada $x \in M$, o isomorfismo $f'(x): T_xM \to T_{f(x)}N$ preserva a orientação nesses espaços vetoriais orientados. Dizemos que f *inverte orientação* (ou é *negativo*) quando, *para todo* $x \in M$, o isomorfismo linear $f'(x)$ inverte orientação. Se M é desconexa, é bem possível que um difeomorfismo local $f: M \to N$ não seja positivo nem negativo, isto é, que preserve a orientação em algumas componentes e inverta em outras. Mostraremos, a seguir, que isto não pode ocorrer quando M é conexa.

Teorema 16. *Sejam M, N superfícies orientadas de mesma dimensão e $f: M \to N$ um difeomorfismo local de classe C^1. O conjunto dos pontos $x \in M$ nos quais a derivada $f'(x): T_xM \to T_{f(x)}N$ é positiva (respect. negativa) é um subconjunto aberto de M.*

Demonstração: Dados $x \in M$ e $y = f(x) \in N$, sejam $\varphi: U_0 \to U \subset M$ e $\psi: V_0 \to V \subset N$ parametrizações positivas com $x \in U$, $y \in V$ e $f(U) \subset V$. Então $f'(x): T_xM \to T_yN$ é positivo se, e somente se, o determinante jacobiano de $(\psi^{-1} \circ f \circ \varphi)'(\varphi^{-1}(x)): \mathbb{R}^m \to \mathbb{R}^m$ é positivo. Como esse determinante é uma função contínua de x, o teorema fica provado.

Corolário 1. *Sejam M, N superfícies orientadas da mesma dimensão. Se M é conexa, um difeomorfismo local $f: M \to N$, de classe C^1, é positivo ou negativo.*

Corolário 2. *Numa superfície conexa orientável, só há duas orientações possíveis.*

Com efeito, sejam $\mathcal{O}$ e $\mathcal{O}'$ orientações em M. A aplicação identidade $f: (M, \mathcal{O}) \to (M, \mathcal{O}')$ é um difeomorfismo local. Logo, ou f é positivo (e então $\mathcal{O} = \mathcal{O}'$) ou é negativo (e, neste caso $\mathcal{O}' = -\mathcal{O}$).

Exemplo 41. A aplicação $\alpha: S^m \to S^m$, $\alpha(x) = -x$, é um difeomorfismo de S^m, com $\alpha^{-1} = \alpha$, chamado a *aplicação antípoda*. Examinemos se α preserva ou inverte a orientação de S^m. Dado $x \in S^m$, escrevamos $E_x = T_x(S^m)$. Então $E_x = E_{-x}$, já que são complementos ortogonais

de x e $-x$. A orientação de S^m, definida pelo campo de vetores normais $v(x) = x$, em conformidade com o Teorema 15, faz com que uma base $\{w_1, \ldots, w_m\} \subset E_x$ seja positiva se, e somente se, $\det(w_1, \ldots, w_m, x) > 0$, onde $(w_1, \ldots, w_m, x)$ é a matriz $(m+1) \times (m+1)$ cujas colunas estão aí indicadas. Segue-se imediatamente que embora os espaços vetoriais E_x e E_{-x} sejam iguais, suas orientações são opostas, isto é, $\mathcal{O}_{-x} = -\mathcal{O}_x$. A derivada $\alpha'(x)\colon E_x \to E_{-x}$ é a multiplicação por -1, logo seu determinante é igual a $(-1)^m$. Portanto, em relação às orientações $\mathcal{O}_x$ e $\mathcal{O}_{-x}$ adotadas nesses espaços, $\alpha'(x)$ é um isomorfismo positivo se, e somente se, m é ímpar. Assim, a aplicação antípoda $\alpha\colon S^m \to S^m$ preserva a orientação de S^m quando m é ímpar e a inverte quando m é par.

Resulta do Exemplo acima uma nova maneira de mostrar que o plano projetivo P^2 não é orientável. Com efeito, suponhamos, por absurdo, que P^2 pudesse ser munido de uma orientação. Então o difeomorfismo local $f\colon S^2 \to P^2$, definido no Exemplo 40, seria positivo ou negativo, em virtude do Corolário 1 acima. Ora, considerando o difeomorfismo negativo $\alpha\colon S^2 \to S^2$, vemos que o composto $f \circ \alpha$ é positivo se f é negativo e vice-versa. Mas $f \circ \alpha = f$, o que nos leva a uma contradição.

15 O método dos multiplicadores de Lagrange

Agora que já temos a noção de superfície de codimensão maior do que 1 e conhecemos o Teorema da Aplicação Implícita, podemos estender o estudo feito no parágrafo 10 do Capítulo III a funções definidas em superfícies descritas por meio de várias equações simultâneas: $f_1(x) = c_1, \ldots, f_n(x) = c_n$. Então, em vez de um, teremos n multiplicadores de Lagrange. Retornemos às definições básicas.

Sabemos que uma função $f\colon M \to \mathbb{R}$, definida numa superfície M, diz-se diferenciável quando, para toda parametrização $\varphi\colon V_0 \to V$ em M, a composta $f \circ \varphi\colon V_0 \to \mathbb{R}$ é uma função diferenciável. Analogamente se define função k vezes diferenciável e função de classe C^k, conceitos que têm sentido quando M é uma superfície de classe C^k. Um exemplo freqüente ocorre quando f é a restrição de uma função $\overline{f}\colon U \to \mathbb{R}$, definida num aberto do espaço euclidiano que contém a superfície M. Neste caso, a diferenciabilidade de $\overline{f}$ implica a de f em virtude da Regra da Cadeia, pois $f \circ \varphi = \overline{f} \circ \varphi$ para toda parametrização de M.

Sabemos também que a diferencial de uma função $f\colon M \to \mathbb{R}$ no ponto $p \in M$ é o funcional linear $df(p)\colon T_pM \to \mathbb{R}$, que pode ser de-

finido assim: dado o vetor $v \in T_pM$, seja $\lambda\colon (-\varepsilon, \varepsilon) \to M$ um caminho diferenciável no ponto 0, com $\lambda(0) = p$ e $\lambda'(0) = v$. Então $df(p) \cdot v = (f \circ \lambda)'(0)$. Em particular, quando f é a restrição a M de uma função $\overline{f}\colon U \to \mathbb{R}$, diferenciável num aberto contendo M e contido no espaço euclidiano onde M está situada, vemos que $df(p) \cdot v = d\overline{f}(p) \cdot v$ para todo $v \in T_pM$. Ou seja, $df(p)$ é a restrição de $d\overline{f}(p)$ ao subespaço vetorial T_pM. Considerando o vetor gradiente de $\overline{f}$, temos ainda $df(p) \cdot v = \langle \operatorname{grad} \overline{f}(p), v\rangle$.

Um *ponto crítico* da função $f\colon M \to \mathbb{R}$ é um ponto $p \in M$ tal que $df(p) = 0$. Quando $f = \overline{f}|M$ é a restrição de $\overline{f}\colon U \to \mathbb{R}$, definida num aberto do espaço euclidiano que contém M, dizer que $p \in M$ é um ponto crítico de f significa o mesmo que afirmar que $\operatorname{grad} \overline{f}(p)$ é um vetor normal a M no ponto p.

Entre os pontos críticos de uma função $f\colon M \to \mathbb{R}$ acham-se os pontos de máximo local e mínimo local de f pois se $p \in M$ é um desses pontos então, para todo caminho $\lambda\colon (-\varepsilon, \varepsilon) \to M$, diferenciável no ponto 0, com $\lambda(0) = p$, a função real $f \circ \lambda\colon (-\varepsilon, \varepsilon) \to \mathbb{R}$ tem um máximo ou um mínimo local no ponto 0, logo $(f \circ \lambda)'(0) = 0$. Assim $df(p) \cdot v = 0$ para todo $v \in T_pM$.

Exemplo 42. Consideremos uma superfície $M \subset \mathbb{R}^n$ e um ponto $a \in \mathbb{R}^n$. Suponhamos que exista, entre os pontos de M, um que esteja mais próximo de a (ou mais afastado) que os demais. Seja p esse ponto. Então a função $f\colon M \to \mathbb{R}$, definida por $f(x) = |x - a|^2$ (norma euclidiana), tem p como ponto crítico. [Com efeito, se a distância é máxima ou mínima, seu quadrado também o é. A vantagem de usar o quadrado é que as fórmulas ficam mais simples. Poderíamos também dizer que $g(x) = |x - a|$ não é diferenciável no ponto $x = a$ mas esta dificuldade seria evitada simplesmente tomando $a \notin M$.] Neste caso, f é obviamente a restrição a M da função $x \mapsto |x - a|^2$, de classe C^∞ em todo o espaço euclidiano $\mathbb{R}^n$. O gradiente desta função num ponto x é $2(x - a)$. Vemos portanto que se $p \in M$ é o ponto de M mais próximo (ou mais afastado) do ponto a, então $2(p - a)$ é normal a M no ponto p, ou seja, $(p - a) \perp T_pM$.

Suponhamos agora que a superfície $M \subset \mathbb{R}^{m+n}$ seja dada como imagem inversa $M = \varphi^{-1}(c)$ de um valor regular $c = (c_1, \ldots, c_n)$ da aplicação $\varphi\colon U \to \mathbb{R}^n$, de classe C^1, definida no aberto $U \subset \mathbb{R}^{m+n}$. Se as funções-coordenada de φ são $\varphi_1, \ldots, \varphi_n$, então a condição de c ser valor regular de φ significa que, para todo $x \in M$, os veto-

res $\operatorname{grad}\varphi_1(x),\ldots,\operatorname{grad}\varphi_n(x)$ são linearmente independentes e portanto constituem uma base do espaço vetorial $T_xM^{\perp}$, normal a M no ponto x. Neste caso, dada uma função diferenciável $f\colon U\to\mathbb{R}$, definida no mesmo aberto U, o ponto $p\in M$ é um ponto crítico da restrição $f|M$ se, e somente se, o vetor $\operatorname{grad} f(p)$ é uma combinação linear dos vetores $\operatorname{grad}\varphi_i(p)$, ou seja, existem números reais $\lambda_1,\ldots,\lambda_n$ tais que

$$\operatorname{grad} f(p)=\lambda_1\cdot\operatorname{grad}\varphi_1(p)+\cdots+\lambda_n\cdot\operatorname{grad}\varphi_n(p).$$

Os números $\lambda_1,\ldots,\lambda_n$ chamam-se *multiplicadores de Lagrange*. Observe que os vetores $\operatorname{grad} f$ e $\operatorname{grad}\varphi_i$, bem como o ponto p, pertencem ao espaço $\mathbb{R}^{m+n}$. Logo, a equação vetorial acima, juntamente com as equações escalares $\varphi_1(p)=c_1,\ldots,\varphi_n(p)=c_n$, constituem um sistema de $m+2n$ equações com $m+2n$ incógnitas (a saber: as $m+n$ coordenadas de p mais os n multiplicadores λ_i). A solução completa desse sistema fornecerá todos os pontos críticos da restrição $f|M$, ou, como às vezes se diz, os pontos críticos de f sujeitos aos n *vínculos* $\varphi_1(x)=c_1,\ldots,\varphi_n(x)=c_n$.

Para registro e posterior referência, enunciemos o

Teorema dos Multiplicadores de Lagrange. *Sejam $f\colon U\to\mathbb{R}$ uma função diferenciável num aberto $U\subset\mathbb{R}^{m+n}$ e $M=\varphi^{-1}(c)$ uma superfície contida em U, imagem inversa do valor regular $c\in\mathbb{R}^n$ por uma aplicação $\varphi\colon U\to\mathbb{R}^n$, de classe C^1. Um ponto $p\in M$ é ponto crítico da restrição $f|M$ se, e somente se, existem n números $\lambda_1,\ldots,\lambda_n$ tais que* $\operatorname{grad} f(p)=\lambda_1\cdot\operatorname{grad}\varphi_1(c)+\cdots+\lambda_n\operatorname{grad}\varphi_n(c)$, *onde $\varphi_1,\ldots,\varphi_n\colon U\to\mathbb{R}$ são as funções-coordenada de φ.*

Exemplo. Seja $A\colon\mathbb{R}^m\to\mathbb{R}^n$ uma transformação linear arbitrária. Sua *adjunta* é a transformação linear $A^*\colon\mathbb{R}^n\to\mathbb{R}^m$, que associa a cada vetor $y\in\mathbb{R}^m$ o vetor $A^*\cdot y\in\mathbb{R}^n$ tal que $\langle A\cdot x,y\rangle=\langle x,A^*\cdot y\rangle$, seja qual for $x\in\mathbb{R}^m$. Consideremos a função bilinear $f\colon\mathbb{R}^m\times\mathbb{R}^n\to\mathbb{R}$ definida por $f(x,y)=\langle A\cdot x,y\rangle=\Sigma\,a_{ij}y_ix_j$ $(1\le i\le n, 1\le j\le m)$, e a superfície $M=S^{m-1}\times S^{n-1}=\varphi^{-1}(1,1)$, onde $\varphi\colon\mathbb{R}^m\times\mathbb{R}^n\to\mathbb{R}^2$ é dada por $\varphi(x,y)=(|x|^2,|y|^2)$. Vejamos quais são os pontos críticos da restrição $f|M$. Para isto, usaremos o método dos multiplicadores de Lagrange.

Para todo $(x,y)\in\mathbb{R}^m\times\mathbb{R}^n$, temos $\operatorname{grad} f(x,y)=(A^*\cdot y,A\cdot x)\in\mathbb{R}^m\times\mathbb{R}^n$. Por outro lado, temos $\varphi=(\varphi_1,\varphi_2)$, onde $\operatorname{grad}\varphi_1(x,y)=(2x,0)$ e $\operatorname{grad}\varphi_2(x,y)=(0,2y)$. Tomando, por conveniência, $\lambda/2$ e $\mu/2$ como

multiplicadores de Lagrange, um ponto $p = (x, y) \in M$ é um ponto crítico de $f|M$ se, e somente se,

$$\operatorname{grad} f(x, y) = \frac{\lambda}{2} \operatorname{grad} \varphi_1(x, y) + \frac{\mu}{2} \operatorname{grad} \varphi_2(x, y),$$

ou seja: $(A^* \cdot y, A \cdot x) = (\lambda \cdot x, \mu \cdot y)$.

Daí obtemos $A \cdot x = \mu \cdot y$, $A^* \cdot y = \lambda \cdot x$, e conseqüentemente $\mu = \langle \mu \cdot y, y \rangle = \langle A \cdot x, y \rangle = \langle x, A^* \cdot y \rangle = \langle x, \lambda \cdot x \rangle = \lambda$. Portanto, os pontos críticos de $f|M$ são os pontos $p = (x, y) \in S^{m-1} \times S^{n-1}$ tais que $A \cdot x = \lambda \cdot y$ e $A^* \cdot y = \lambda \cdot x$ para um certo $\lambda \in \mathbb{R}$. Note que, neste caso, $\lambda = f(x, y)$ e, além disso, $z \perp x \Rightarrow A \cdot z \perp y$. Portanto, se chamarmos de $E = \{z \in \mathbb{R}^m; z \perp x\}$ o complemento ortogonal de X em $\mathbb{R}^m$ e $F = \{w \in \mathbb{R}^n; w \perp y\}$ o complemento ortogonal de y em $\mathbb{R}^n$, a transformação linear A aplica E em F.

Seja então $p_1 = (u_1, v_1) \in S^{m-1} \times S^{n-1}$ o ponto onde a função f assume seu valor máximo $f(u_1, v_1) = \lambda_1$ em $S^{m-1} \times S^{n-1}$. Temos $A \cdot u_1 = \lambda_1 \cdot v_1$ e $A^* \cdot v_1 = \lambda_1 \cdot u_1$. Como $f(x, -y) = -f(x, y)$, vemos que $\lambda_1 \geq 0$. Se $A \neq 0$ então f não é identicamente nula em M e tem-se $\lambda_1 > 0$.

Em seguida, tomemos $E \subset \mathbb{R}^m$ e $F \subset \mathbb{R}^n$, complementos ortogonais de u_1 e v_1 respectivamente. Sabemos que $A \cdot E \subset F$. Podemos então considerar a restrição $A|E$ como uma transformação linear $A \colon E \to F$, com $\dim E = m - 1$ e $\dim F = n - 1$.

Aplicando a esta transformação o mesmo raciocínio usado para $A \colon \mathbb{R}^m \to \mathbb{R}^n$ e prosseguindo analogamente, chegaremos ao seguinte resultado:

Dada qualquer transformação linear $A \colon \mathbb{R}^m \to \mathbb{R}^n$, com posto k, existem bases ortonormais $\{u_1, \ldots, u_m\}$ em $\mathbb{R}^m$, $\{v_1, \ldots, v_n\}$ em $\mathbb{R}^n$ e números reais positivos $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_k$ tais que $A \cdot u_i = \lambda_i \cdot v_i$, $A^* \cdot v_i = \lambda_i \cdot u_i$ $(i = 1, \ldots, k)$ e $A \cdot u_i = 0$, $A^* \cdot v_i = 0$ se $i > k$.

Em termos de matrizes, isto significa que, dada qualquer matriz A, com n linhas e m colunas, existem matrizes ortogonais $P \in \mathbb{R}^{n^2}$, $Q \in \mathbb{R}^{m^2}$ tais que a matriz $PAQ = \Lambda$ tem a forma $\Lambda = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}$, onde $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_k)$ é uma matriz diagonal com $\lambda_1 \geq \cdots \geq \lambda_k > 0$.

Observação: Os números $\lambda_1^2, \ldots, \lambda_k^2$ são os valores próprios não-nulos de transformação linear simétrica $A^*A \colon \mathbb{R}^m \to \mathbb{R}^m$ (ou, equivalentemente, de $AA^* \colon \mathbb{R}^n \to \mathbb{R}^n$).

Exercícios

§1.

1.1. Sejam $\alpha > 1$ e $c \in \mathbb{R}$. Se $f\colon U \to \mathbb{R}^n$, definida no aberto $U \subset \mathbb{R}^m$, cumpre a condição $|f(x) - f(y)| \leq c|x - y|^\alpha$ para quaisquer $x, y \in U$ então f é constante em cada componente de U.

1.2. Sejam $U \subset \mathbb{R}^m$ aberto e $f, g\colon U \to \mathbb{R}^n$ diferenciáveis no ponto $a \in U$, com $f(a) = g(a)$. A fim de que $f'(a) = g'(a)$, é necessário e suficiente que $\lim\limits_{v \to 0} \dfrac{f(a+v) - g(a+v)}{|v|} = 0$.

1.3. Sejam $V \subset U \subset \mathbb{R}^m$ abertos e $\delta > 0$ tais que $x \in V$, $|h| < \delta \Rightarrow x+h \in U$. Seja $B = B(0;\delta)$. Se $f\colon U \to \mathbb{R}^n$ é diferenciável então $\varphi\colon V \times B \to \mathbb{R}^n$, definida por $\varphi(x,h) = f(x+h)$, é diferenciável, sendo $\varphi'(x_0, h_0)\colon \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^n$ dada por $\varphi'(x_0, h_0) \cdot (u, v) = f'(x_0 + h_0) \cdot (u + v)$.

1.4. Seja $U \subset \mathbb{R}^m$ aberto. A fim de que uma aplicação $f\colon U \to \mathbb{R}^n$ seja diferenciável no ponto $a \in U$ é necessário e suficiente que exista, para cada $h \in \mathbb{R}^m$ com $a+h \in U$, uma transformação linear $A(h)\colon \mathbb{R}^m \to \mathbb{R}^n$ tal que $f(a+h) - f(a) = A(h) \cdot h$ e $h \mapsto A(h)$ seja contínua no ponto $h = 0$.

1.5. Dado $U \subset \mathbb{R}^m$ aberto, seja $f\colon U \to \mathbb{R}^n$ diferenciável no ponto $a \in U$. Prove que se $\lim v_k = v$ em $\mathbb{R}^m$ e $\lim t_k = 0$ em $\mathbb{R}$ então $\lim\limits_{k \to \infty} \dfrac{f(a + t_k v_k) - f(a)}{t_k} = f'(a) \cdot v$.

1.6. Seja $f\colon U \to \mathbb{R}^n$ diferenciável no aberto $U \subset \mathbb{R}^m$. Se, para algum $b \in \mathbb{R}^n$, o conjunto $f^{-1}(b)$ possui um ponto de acumulação $a \in U$ então $f'(a)\colon \mathbb{R}^m \to \mathbb{R}^n$ não é injetiva.

1.7. Dada $f\colon S^m \to \mathbb{R}^n$, defina a *extensão radial* de f como a aplicação $F\colon \mathbb{R}^{m+1} \to \mathbb{R}^n$ tal que $F(0) = 0$ e $F(x) = |x| \cdot f\left(\dfrac{x}{|x|}\right)$ se $x \neq 0$. Mostre que F é diferenciável na origem $0 \in \mathbb{R}^{m+1}$ se, e somente se, f é (a restrição de uma aplicação) linear.

1.8. Seja $U \subset \mathbb{R}^m$ aberto. Uma aplicação $A\colon U \to \mathcal{L}(\mathbb{R}^n; \mathbb{R}^p)$ é diferenciável no ponto $a \in U$ se, e somente se, para cada vetor $w \in \mathbb{R}^n$, a aplicação $\varphi_w\colon U \to \mathbb{R}^p$, definida por $\varphi_w(x) = A(x) \cdot w$, é diferenciável no ponto a. No caso afirmativo, tem-se $\varphi'_w(a) \cdot v = (A'(a) \cdot v) \cdot w$.

1.9. Dada $f\colon \mathbb{R}^n \to \mathbb{R}^p$, enuncie e demonstre um teorema que traduza a igualdade $f'(x,y) = \dfrac{\partial f}{\partial x}\, dx + \dfrac{\partial f}{\partial y}\, dy$.

1.10. Seja $f\colon U \to \mathbb{R}^p$ duas vezes diferenciável no aberto $U \subset \mathbb{R}^m \times \mathbb{R}^n$. Defina as derivadas mistas $\dfrac{\partial^2 f}{\partial x \partial y}$, $\dfrac{\partial^2 f}{\partial y \partial x}$ e estabeleça a relação que existe entre elas.

1.11. Seja $f\colon \mathbb{R}^m \to \mathbb{R}^m$ diferenciável, com $f(0) = 0$. Se a transformação linear $f'(0)$ não tem valor próprio 1 então existe uma vizinhança V de 0 em $\mathbb{R}^m$ tal que $f(x) \neq x$ para todo $x \in V - \{0\}$.

1.12. Seja $f\colon U \to \mathbb{R}^n$ duas vezes diferenciável no ponto $a \in U$, U aberto em $\mathbb{R}^m$. Se $\lim t_k = 0$ em $\mathbb{R}$ e $\lim v_k = v$, $\lim w_k = w$ em $\mathbb{R}^m$ então

$$f''(a)\cdot v\cdot w =$$

$$= \lim_{k\to\infty} \frac{f(a+t_k(v_k+w_k)) - f(a+t_kv_k) - f(a+t_kw_k) + f(a)}{t_k^2}.$$

Em particular, $f''(a)\cdot v\cdot w = \lim\limits_{t\to 0} \frac{1}{t^2}\,[f(a+t(v+w)) - f(a+tv) - f(a+tw) + f(a)]$. Conclua daí uma nova demonstração do Teorema de Schwarz.

1.13. Sejam $U \subset \mathbb{R}^m$ um aberto simplesmente conexo e $A\colon U \to \mathcal{L}(\mathbb{R}^m;\mathbb{R}^n)$ uma aplicação diferenciável. A fim de que exista $f\colon U \to \mathbb{R}^n$ duas vezes diferenciável, tal que $f'(x) = A(x)$ para todo $x \in U$, é necessário e suficiente que $[A'(x)\cdot v]\cdot w = [A'(x)\cdot w]\cdot v$ para $v, w \in \mathbb{R}^m$ quaisquer.

§3.

3.1. Sejam $f\colon U \to \mathbb{R}^n$ Lipschitziana no aberto $U \subset \mathbb{R}^m$, com $a \in U$, e $g\colon V \to \mathbb{R}^p$ diferenciável no aberto $V \subset \mathbb{R}^n$, com $f(U) \subset V$ e $b = f(a)$. Se $g'(b) = 0$ então $g\circ f\colon U \to \mathbb{R}^p$ é diferenciável no ponto a, com $(g\circ f)'(a) = 0$.

3.2. Seja $f\colon U \to \mathbb{R}^n$ Lipschitziana no aberto $U \subset \mathbb{R}^m$. Dado $a \in U$, suponha que, para todo $v \in \mathbb{R}^m$, exista a derivada direcional $\frac{\partial f}{\partial v}(a)$ e dependa linearmente de v. Prove que, para todo caminho $g\colon (-\varepsilon,\varepsilon) \to U$, com $g(0) = a$, diferenciável no ponto $t = 0$, existe o vetor-velocidade $(f\circ g)'(0)$. Conclua que f é diferenciável no ponto a.

3.3. Seja $U \subset \mathbb{R}^m$ um aberto como no segundo exercício do §3, Capítulo III. Defina aplicação positivamente homogênea $f\colon U \to \mathbb{R}^n$, de grau k. Prove que a relação de Euler $f'(x)\cdot x = k\cdot f(x)$ é necessária e suficiente para que uma aplicação diferenciável $f\colon U \to \mathbb{R}^n$ seja positivamente homogênea de grau k. Enuncie e prove também uma versão do 11º exercício do §8, Capítulo III, para aplicações homogêneas com valores em $\mathbb{R}^n$.

3.4. Sejam $f\colon U \to \mathbb{R}^n$ e $g\colon V \to \mathbb{R}^m$ diferenciáveis nos abertos $U \subset \mathbb{R}^m$ e $V \subset \mathbb{R}^n$, com $g(f(x)) = x$ para todo $x \in U$. Se $y = f(x)$, prove que as transformações lineares $f'(x)\colon \mathbb{R}^m \to \mathbb{R}^n$ e $g'(y)\colon \mathbb{R}^n \to \mathbb{R}^m$ têm o mesmo posto.

3.5. Sejam $f\colon U \to \mathbb{R}^m$ diferenciável no aberto $U \subset \mathbb{R}^m$ e $\varphi\colon \mathbb{R}^m \to \mathbb{R}$ de classe C^1, com $\varphi(f(x)) = 0$ para todo $x \in U$. Dado $a \in U$, se $\operatorname{grad}\varphi(b) \neq 0$, $b = f(a)$, então $\det .f'(a) = 0$.

3.6. Seja $f\colon U \to \mathbb{R}^m$ diferenciável no aberto $U \subset \mathbb{R}^m$. Se $|f(x)|$ é constante quando x varia em U então o determinante jacobiano de f é identicamente nulo.

3.7. Enuncie precisamente e demonstre as afirmações abaixo, encontradas num livro de Cálculo:

a) "Substituindo as coordenadas cartesianas x, y pelas coordenadas polares r, θ, com $x = r\cos\theta$ e $y = r\operatorname{sen}\theta$, na função $w = f(x,y)$ tem-se

$$\frac{\partial w}{\partial r} = \frac{\partial f}{\partial x}\cos\theta + \frac{\partial f}{\partial y}\operatorname{sen}\theta \text{ e } \frac{1}{r}\frac{\partial w}{\partial \theta} = -\frac{\partial f}{\partial x}\operatorname{sen}\theta + \frac{\partial f}{\partial y}\cos\theta.''$$

b) "Sejam $w = w(u, v)$, $u = u(x, y, z)$ e $v = v(x, y, z)$ funções diferenciáveis. Considerando-se w como função de x, y, z, seu gradiente em cada ponto (x_0, y_0, z_0) está no mesmo plano que os gradientes de u e v neste ponto."

3.8. Sejam $U = (0, +\infty) \times (0, 2\pi) \times \mathbb{R}$ e $V = \mathbb{R}^3$ menos o semi-plano $y = 0$, $x \geq 0$. Mostre que $\varphi\colon U \to V$, definida por $\varphi(r, \theta, z) = (r \cdot \cos\theta, r \cdot \operatorname{sen}\theta, z)$ é um difeomorfismo C^∞. (Se $q = \varphi(r, \theta, z)$, os números r, θ e z são chamados as "coordenadas cilíndricas "de q.) Dada $f\colon V \to \mathbb{R}$ diferenciável, explique o significado e demonstre a seguinte fórmula para o gradiente de f em coordenadas cilíndricas:

$$\operatorname{grad} f = \frac{\partial f}{\partial r} \cdot u_r + \frac{1}{r}\frac{\partial f}{\partial \theta} u_\theta + \frac{\partial f}{\partial z} u_z\,,$$

onde u_r, u_θ e u_z são os vetores unitários tangentes às curvas r, θ e z em V.

3.9. a) Sejam $U \subset \mathbb{R}^m$, $V \subset \mathbb{R}^n$ abertos e $f\colon U \to \mathbb{R}^n$, $g\colon V \to \mathbb{R}$ diferenciáveis, com $f(U) \subset V$. Para todo $x \in U$, ponto $y = f(x)$ tem-se

$$\operatorname{grad}(g \circ f)(x) = [f'(x)]^* \cdot \operatorname{grad} g(y).$$

b) Interprete a igualdade $\operatorname{grad}_x z = \sum_i \dfrac{\partial z}{\partial y_i} \cdot \operatorname{grad}_x y_i$, escrita na notação clássica, e identifique-a com o resultado anterior.

3.10. Sejam $U \subset \mathbb{R}^m$ aberto e $A\colon U \to \mathcal{L}(\mathbb{R}^n; \mathbb{R}^p)$ diferenciável. Defina $g\colon U \times \mathbb{R}^n \to \mathbb{R}^p$ pondo $g(x, v) = A(x) \cdot v$. Prove que g é diferenciável e que $g'(x, y)\colon \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^p$ é dada por $g'(x, v) \cdot (h, k) = (A'(x) \cdot h) \cdot v + A(x) \cdot k$.

3.11. Seja $f\colon U \to \mathbb{R}^n - \{0\}$ diferenciável no aberto conexo $U \subset \mathbb{R}^m$. A fim de que seja $|f(x)| =$ constante, é necessário e suficiente que $f'(x) \cdot h$ seja perpendicular a $f(x)$, para todo $x \in U$ e todo $h \in \mathbb{R}^m$.

3.12. Enuncie precisamente e demonstre a seguinte fórmula para a derivada da composta de duas aplicações em abertos do espaço euclidiano, com $y = f(x)$:

$$(g \circ f)''(x) = g''(y) \cdot f'(x)^2 + g'(y) \cdot f''(x).$$

3.13. Sejam $f, g, h\colon U \to \mathbb{R}^m$ diferenciáveis no aberto $U \subset \mathbb{R}^m$ contendo a origem, com $f(0) = g(0) = 0$. Se $h \circ f = g \circ h$ e $h'(0)\colon \mathbb{R}^m \to \mathbb{R}^m$ for um isomorfismo, prove que $f'(0)$ e $g'(h(0))$ possuem os mesmos valores próprios.

3.14. Seja $f\colon U \to \mathbb{R}^n$ diferenciável no aberto $U \subset \mathbb{R}^m$, com $f(x) \neq 0$ para todo $x \in U$.
a) Defina $\varphi\colon U \to \mathbb{R}$ pondo $\varphi(x) = 1/|f(x)|$ (norma euclidiana). Para todo $x \in U$ e todo $v \in \mathbb{R}^m$, determine $\varphi'(x) \cdot v$.
b) Seja $g\colon U \to \mathbb{R}^k$ diferenciável. Defina $\xi\colon U \to \mathbb{R}^k$ pondo $\xi(x) = g(x)/|f(x)|$. Calcule $\xi'(x) \cdot v$ para todo $x \in U$ e todo $v \in \mathbb{R}^m$.

3.15. Sejam $U \subset \mathbb{R}^m$ e $V \subset \mathbb{R}^n$ abertos, $f\colon U \to \mathbb{R}^n$ uma aplicação i vezes diferenciável no ponto $x \in U$, com $f(U) \subset V$, e $g\colon V \to \mathbb{R}^p$ uma aplicação i vezes diferenciável no ponto $y = f(x)$. Para cada partição $i = i_1 + \cdots + i_k$

(de i como soma de números naturais) existe um inteiro $n(i_1, \ldots, i_k)$ tal que a i-ésima derivada da aplicação composta $g \circ f \colon U \to \mathbb{R}^p$ tem a expressão

$$(g \circ f)^{(i)} = \sum_{k=1}^{i} n(i_1, \ldots, i_k) g^{(k)} \circ f \cdot (f^{(i_1)}, \ldots, f^{(i_k)}).$$

3.16. Sejam $f, g \colon V \to \mathbb{R}^n$ diferenciáveis no aberto $V \subset \mathbb{R}^m$, δ um número real positivo e $X \subset V$. Escreva "$|f - g|_1 \leq \delta$ em X"para significar que $|f(x) - g(x)| \leq \delta$ e $|f'(x) - g'(x)| \leq \delta$ para todo $x \in X$. Prove que se $\varphi \colon U \to V$ é de classe C^1 no aberto $U \subset \mathbb{R}^k$ então, dados $K \subset U$ compacto e $\eta > 0$, existe $\delta > 0$ tal que $|f - g|_1 < \delta$ em $\varphi(K)$ implica $|f \circ \varphi - g \circ \varphi|_1 < \eta$ em K.

§5.

5.1. Dado o aberto conexo $U \subset \mathbb{R}^m$, defina a *distância geodésica* $d_U(x, y)$ entre os pontos $x, y \in U$ como o ínfimo dos comprimentos dos caminhos (contínuos) e retificáveis contidos em U, que começam em x e terminam em y. Mostre que, sem perda de generalidade, na definição de $d_U(x, y)$ basta considerar caminhos poligonais. Mostre que se $f \colon U \to \mathbb{R}^n$ é diferenciável e $|f'| \leq M$ para todo $x \in U$ então $|f(y) - f(x)| \leq M \cdot d_U(x, y)$ para quaisquer $x, y \in U$. Se U é convexo e $V \subset \mathbb{R}^n$ é um aberto conexo contendo $f(U)$, mostre que $x, y \in U \Rightarrow d_V(f(x), f(y)) \leq M \cdot |x - y|$ desde que $|f'(x)| \leq M$ para todo $x \in U$.

5.2. Seja $f \colon U \to \mathbb{R}^n$ de classe C^1 num aberto convexo $U \subset \mathbb{R}^m$, com $0 \in U$ e $f(0) = 0$. Se $|f'(x)| \leq |x|$ para todo $x \in U$ então $|f(x)| \leq \frac{1}{2} |x|^2$ seja qual for $x \in U$. Conclua que se $f(0) = 0$ e $f'(0) = 0$, com $f \in C^2$, então $\left| \frac{\partial^2 f}{\partial u \partial v}(x) \right| \leq |u|\,|v|$ para $x \in U$, $u, v \in \mathbb{R}^m$ quaisquer implica ainda $|f(x)| \leq \frac{1}{2} |x|^2$. Generalize.

5.3. Sejam $U \subset \mathbb{R}^m$ aberto, $[a, b] \subset U$, $f \colon U \to \mathbb{R}^n$ contínua em $[a, b]$ e diferenciável em (a, b). Para cada $y \in \mathbb{R}^n$ existe $c_y \in (a, b)$ tal que $\langle f(b) - f(a), y \rangle = \langle f'(c_y) \cdot (b - a), y \rangle$.

5.4. Seja $U \subset \mathbb{R}^m$ convexo. Dada $f \colon U \to \mathbb{R}^n$ diferenciável, considere as seguintes afirmações:

a) $|f'(x)| \leq c$ para todo $x \in U$;

b) $|f(x) - f(y)| \leq c|x - y|$ para quaisquer $x, y \in U$;

c) f é uniformemente contínua:

d) Para todo $x_0 \in \overline{U}$, existe $\lim_{x \to x_0} f(x)$;

e) Se U é limitado então $f(U)$ é limitado.

Mostre que $a \Leftrightarrow b \Rightarrow c \Rightarrow d \Rightarrow e$ mas as demais implicações são todas falsas.

5.5. Seja $U \subset \mathbb{R}^m$ aberto conexo. Se $f \colon U \to \mathbb{R}^n$ é diferenciável e $f'(x) = T$ (constante) para todo $x \in U$ então existe $a \in \mathbb{R}^m$ tal que $f'(x) = T \cdot x + a$. Mais geralmente, se $f^{(k+1)}(x) = 0$ para todo $x \in U$ então f é um "polinômio de grau k"(soma de formas multilineares de grau $\leq k$, restritas à diagonal).

5.6. Seja $f: U \to \mathbb{R}^n$ diferenciável no aberto convexo $U \subset \mathbb{R}^m$. Mostre que

$$\sup_{x \neq y} \frac{|f(x) - f(y)|}{|x - y|} = \sup_{z \in U} |f'(z)|.$$

5.7. Se $f: \mathbb{R}^m \to \mathbb{R}^n$ é diferenciável, com $\lim\limits_{|x| \to \infty} f'(x) \cdot x = 0$ então a aplicação $g: \mathbb{R}^m \to \mathbb{R}^n$, definida por $g(x) = f(2x) - f(x)$, é limitada.

5.8. Seja $f: U \to \mathbb{R}^n$ diferenciável no aberto $U \subset \mathbb{R}^m$. Se $f': U \to \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$ é contínua e $K \subset U$ é compacto então existe $a > 0$ tal que $x, y \in K \Rightarrow |f(x) - f(y)| \leq a|x - y|$. O mesmo resultado vale se supusermos f' limitada em U, em vez de contínua.

5.9. Seja $f: U \to \mathbb{R}^n$ diferenciável no aberto $U \subset \mathbb{R}^m$. A fim de que a derivada $f': U \to \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$ seja uniformemente contínua, é necessário e suficiente que, para todo $\varepsilon > 0$ exista $\delta > 0$ tal que $|h| < \delta$, $\quad [x, x + h] \subset U \Rightarrow |f(x + h) - f(x) - f'(x) \cdot h| \leq \varepsilon |h|$.

§6.

6.1. Seja $f: U \to \mathcal{L}(\mathbb{R}^n; \mathbb{R}^n)$ uma aplicação de classe C^∞ num aberto $U \subset \mathbb{R}^m$. Para $k = 1, 2, 3, \dots$, e $x \in U$, seja $f^k(x) = f(x) \cdot f(x) \cdot \dots \cdot f(x)$ (k fatores). Defina $g: U \to \mathcal{L}(\mathbb{R}^n; \mathbb{R}^n)$ pondo, para cada $x \in U$, $\quad g(x) = \sum\limits_{k=1}^{\infty} \frac{1}{k!} f^k(x)$. Prove que g é de classe C^∞.

6.2. Seja $U \subset \mathbb{R}^m$ aberto e conexo. Suponha que uma seqüência de aplicações $f_k: U \to \mathbb{R}^n$, de classe C^∞, convirja num ponto $a \in U$ e que, para cada $i = 1, 2, \dots$, a seqüência das derivadas $(f_k^{(i)})_{k \in \mathbb{N}}$, convirja de modo localmente uniforme em U. Conclua que (f_k) converge de modo localmente uniforme em U para uma aplicação $f: U \to \mathbb{R}^n$, de classe C^∞ e $\lim\limits_{k \to \infty} f_k^{(i)} = f^{(i)}$ para todo $i = 1, 2, \dots$. Enuncie e prove resultados análogos (1) para aplicações de classe C^r $\quad (0 < r < \infty)$; (2) para séries.

6.3. Para toda matriz $X \in \mathbb{R}^{m^2}$, seja $e^X = \sum\limits_{n=0}^{\infty} \frac{X^n}{n!} \cdot$ Prove que a aplicação $\exp: X \mapsto e^X$ é de classe C^∞, e sua derivada no ponto $X = 0$ é a transformação linear identidade.

6.4. Mostre que existe uma aplicação $\log: U \to \mathbb{R}^{n^2}$, de classe C^∞ numa vizinhança U da transformação identidade $I \in \mathbb{R}^{n^2}$, tal que $e^{\log X} = X$ para todo $X \in U$.

§7.

7.1. Seja $f: U \to \mathbb{R}^n$ fortemente diferenciável no aberto $U \subset \mathbb{R}^m$. Se a derivada $f'(x): \mathbb{R}^m \to \mathbb{R}^n$ for injetiva em todos os pontos x pertencentes ao compacto $K \subset U$, prove que existem números reais $c > 0$ e $\delta > 0$ tais que $|f(x) - f(y)| \geq c|x - y|$, quaisquer que sejam $x \in K$, $\quad y \in U$ com $|x - y| < \delta$.

§8.

8.1. Seja $I \subset \mathbb{R}$ um intervalo aberto. Uma função diferenciável $f\colon I \to \mathbb{R}$ é um difeomorfismo sobre $f(I)$ se, e somente se $f'(x) \cdot f'(y) > 0$ para quaisquer $x, y \in I$.

8.2. Seja $f\colon \mathbb{R} \to \mathbb{R}$ uma função de classe C^1 tal que $|f'(t)| \leq k < 1$ para todo $t \in \mathbb{R}$. Defina uma aplicação $\varphi\colon \mathbb{R}^2 \to \mathbb{R}^2$ pondo $\varphi(x,y) = (x + f(y), y + f(x))$. Mostre que φ é um difeomorfismo de $\mathbb{R}^2$ sobre si mesmo.

8.3. Sejam $f, g, h\colon \mathbb{R} \to \mathbb{R}$ diferenciáveis. Defina $F\colon \mathbb{R}^2 \to \mathbb{R}^2$ pondo $F(x,y) = (f(x) \cdot h(y), g(y))$. Suponha que f e g são difeomorfismos de $\mathbb{R}$ sobre $\mathbb{R}$. Mostre que F é um difeomorfismo se, e somente se, $0 \notin h(\mathbb{R})$.

8.4. Sejam $A = \{(x,y) \in \mathbb{R}^2; x > 0, y > 0\}$, $B = \{(x,y) \in \mathbb{R}^2; y > 0\}$ e $C = \{(x,y) \in \mathbb{R}^2; x > 0 \text{ ou } y < 0\}$. Mostre que existem difeomorfismos de classe C^∞, entre A e B e entre A e C. Mostre ainda que *não* existem difeomorfismos $f\colon U \to V$, $g\colon U \to W$, onde $U \supset \overline{A}$, $V \supset \overline{B}$, $W \supset \overline{C}$, tais que $f(\overline{A}) = B$ e $g(\overline{A}) = C$.

8.5. Sejam $f\colon V \to \mathbb{R}$ de classe C^2 e $b \in V$ um ponto crítico de f. Se $\varphi\colon U \to V$ é um difeomorfismo C^2 com $\varphi(a) = b$ então as matrizes Hessianas de f no ponto b e de $f \circ \varphi$ no ponto a têm o mesmo posto.

8.6. Sejam $g\colon [0, +\infty] \to \mathbb{R}$ contínua, com $g(t) > 0$ para todo $t \geq 0$ e $U = \{(x,y) \in \mathbb{R}^2; 0 < x < y\}$. Defina $f\colon U \to \mathbb{R}^2$ pondo

$$f(x,y) = \left(\int_0^{x+y} g(t)\, dt, \int_0^{y-x} g(t)\, dt \right).$$

Mostre que f é um difeomorfismo sobre um aberto de $\mathbb{R}^2$.

8.7. Seja $f\colon U \to \mathbb{R}^m$ de classe C^1 no aberto $U \subset \mathbb{R}^m$. Se as singularidades de f (pontos onde o determinante jacobiano de f se anula) são pontos isolados e $m > 1$ então f é uma aplicação aberta. Deduza daí uma demonstração do Teorema Fundamental da Álgebra.

8.8. Seja $f\colon \mathbb{R}^3 \to \mathbb{R}$ tal que $\dfrac{\partial f}{\partial x} + \dfrac{\partial f}{\partial y} + \dfrac{\partial f}{\partial z} = 0$ em todos os pontos de $\mathbb{R}^3$. Prove que existe um difeomorfismo $g\colon \mathbb{R}^3 \to \mathbb{R}^3$, de classe C^∞, tal que $\dfrac{\partial}{\partial x}(f \circ g) = 0$ em todos os pontos de $\mathbb{R}^3$. [Sugestão: experimente g linear.]

8.9. Seja $f\colon U \to \mathbb{C}$ uma função holomorfa (de classe C^1) no aberto U do plano complexo. Se $f'(z_0) \neq 0$ então z_0 possui uma vizinhança, restrita à qual f tem uma inversa holomorfa.

8.10. Dados $z_0 \in \mathbb{C} - \{0\}$ e $k \in \mathbb{Z}$, existem funções holomorfas $f, g\colon U \to \mathbb{C}$, definidas num aberto $U \ni z_0$, tais que $f(z)^k = e^{g(z)} = z$, para todo $z \in U$. Conclua que, dada $\varphi\colon V \to \mathbb{C} - \{0\}$ de classe C^r num aberto $V \subset \mathbb{R}^m$ e dado $k \in \mathbb{Z}$, existem, para cada $z_0 \in V$, um aberto W, com $z_0 \in W \subset V$ e funções $f, g\colon W \to C$, de classe C^r, tais que $f(x)^k = e^{g(x)} = \varphi(x)$ para todo $x \in W$. Se $m = 2$ e φ for holomorfa, então f e g serão holomorfas.

8.11. Seja $f: U \to \mathbb{R}^m$ diferenciável no conjunto convexo $U \subset \mathbb{R}^m$. Se $\langle f'(x)\cdot v, v\rangle > 0$ para $x \in U$ e $0 \neq v \in \mathbb{R}^m$ quaisquer então f é injetiva. Se $f \in C^1$ então f é um difeomorfismo de U sobre um subconjunto de $\mathbb{R}^m$. Dê um exemplo em que $U = \mathbb{R}^m$ mas f não é sobrejetiva.

8.12. Seja $f: \mathbb{R}^m \to \mathbb{R}^m$ de classe C^1 tal que para $x, v \in \mathbb{R}^m$ quaisquer tem-se $\langle f'(x)\cdot v, v\rangle \geq \alpha|v|^2$, onde $\alpha > 0$ é uma constante. Prove que $|f(x) - f(y)| \geq \alpha|x - y|$ para $x, y \in \mathbb{R}^m$ arbitrários. Conclua que $f(\mathbb{R}^m)$ é fechado, e daí, que f é um difeomorfismo de $\mathbb{R}^m$ sobre si mesmo. (A hipótese "$f \in C^1$"pode ser substituída por "f diferenciável". Veja último exemplo, §10 do Cap. VII.)

8.13. Seja $f: \mathbb{R}^m \to \mathbb{R}^m$ de classe C^1 tal que $f'(x)$ é, para todo $x \in \mathbb{R}^m$, uma isometria (isto é, $|f'(x)\cdot v| = |v|$) na norma euclidiana. Então f é uma isometria (isto é, $|f(x) - f(y)| = |x - y|$). Conclua que existem $T \in \mathcal{L}(\mathbb{R}^m)$ ortogonal e $a \in \mathbb{R}^m$ tais que $f(x) = T\cdot x + a$.

8.14. Seja $f: \mathbb{R}^m \to \mathbb{R}^m$ de classe C^1, com $|f'(x)\cdot v| \geq \alpha|v|$ para $x, v \in \mathbb{R}^m$ quaisquer ($\alpha > 0$ constante). Então f é um difeomorfismo de $\mathbb{R}^m$ sobre si mesmo. Em particular, $|f(x) - f(y)| \geq \alpha|x - y|$, para quaisquer $x, y \in \mathbb{R}^m$.
[Sugestão: prove que, dado qualquer segmento de reta $[f(a), b] \subset \mathbb{R}^m$, existe um caminho $\lambda: [0,1] \to \mathbb{R}^m$ com $\lambda(0) = a$ e $f(\lambda(t)) = (1-t)f(a) + tb$.]
Observação: Este exercício fica bem mais fácil quando se conhece algo sobre "espaços de recobrimento".

8.15. Seja $g: V \to \mathbb{R}^m$ uma aplicação de classe C^2 no aberto $V \subset \mathbb{R}^m$. Dado $b \in V$, suponha que $g'(b): \mathbb{R}^m \to \mathbb{R}^m$ seja um isomorfismo. Prove que existe uma bola aberta B de centro b, tal que a função $\varphi: B \to \mathbb{R}$, definida por $\varphi(y) = |g(y) - g(b)|^2$, é convexa.

8.16. Seja $f: U \to V$ um difeomorfismo de classe C^2 entre abertos $U, V \subset \mathbb{R}^m$. Prove que, para todo ponto $a \in U$ existe $r > 0$ tal que a imagem por f de toda bola de centro a e raio $\leq r$ é um conjunto convexo.

§10.

10.1. Seja $f: I \to \mathbb{R}^2$, dado por $f(t) = (x(t), y(t))$, um caminho diferenciável (respect. de classe C^k). Se $x'(t) \neq 0$ para todo $t \in I$, então $f(I)$ é o gráfico de uma função diferenciável (respect. de classe C^k) $\xi: J \to \mathbb{R}$. Se $I = \mathbb{R}$ e existe uma constante $c > 0$ tal que $|x'(t)| \geq c$ para todo $t \in \mathbb{R}$, então a função ξ é definida para todo $t \in \mathbb{R}$. Seja $f: \mathbb{R}^2 \to \mathbb{R}^3$ definida por $f(s,t) = (x(s,t), y(s,t), z(s,t))$, onde $x(s,t) = e^s\cdot\cos t$, $y(s,t) = e^s\cdot\operatorname{sen} t$ e $z(s,t) = t$. Mostre que f é injetiva, de classe C^∞, que o determinante jacobiano $\frac{\partial x}{\partial s}\frac{\partial y}{\partial t} - \frac{\partial x}{\partial t}\frac{\partial y}{\partial s}$ é diferente de zero em todos os pontos $(s,t) \in \mathbb{R}^2$ mas a imagem $f(\mathbb{R}^2)$ não é o gráfico de uma função real de duas variáveis.

10.2. Seja $f: U \to \mathbb{R}^{m+1}$ de classe C^k $(k \geq 1)$ no aberto $U \subset \mathbb{R}^m$, com

$$f(x) = (f_1(x), \ldots, f_{m+1}(x)) \text{ e } \det\cdot\left(\frac{\partial f_i}{\partial x_j}(x)\right) \neq 0$$

$(1 \leq i, j \leq m)$ para todo $x \in U$. Então cada ponto $x \in U$ possui uma vizinhança $W \subset U$ tal que $f(W)$ é o gráfico de uma função $y_{m+1} = \varphi(y_1, \ldots, y_m)$ de classe C^k.

10.3. Seja $\varphi\colon U \to \mathbb{R}^n$ de classe C^k $(k \geq 1)$ no aberto $U \subset \mathbb{R}^m$. Se $a \in U$ é tal que $\varphi'(a)\colon \mathbb{R}^m \to \mathbb{R}^n$ é injetiva, então existem uma decomposição em soma direta $\mathbb{R}^n = \mathbb{R}^m \oplus \mathbb{R}^{n-m}$ e um aberto V, com $a \in V \subset U$, tais que $\varphi(V)$ é o gráfico de uma aplicação $f\colon W \to \mathbb{R}^{n-m}$, de classe C^k num aberto $W \subset \mathbb{R}^m$.

10.4. Sejam $f, g\colon U \to \mathbb{R}^n$ diferenciáveis no aberto $U \subset \mathbb{R}^m$, δ um número real positivo e $X \subset U$. Escreva "$|f-g|_1 \leq \delta$ em X"para significar que $|f(x) - g(x)| \leq \delta$ e $|f'(x) - g'(x)| \leq \delta$ para todo $x \in X$. Sejam $K \subset U$ compacto e $f\colon U \to \mathbb{R}^n$ de classe C^1, tal que $f|K$ é uma imersão (isto é, a derivada $f'(x)\colon \mathbb{R}^m \to \mathbb{R}^n$ é injetiva para todo $x \in K$). Prove que existe $\delta > 0$ tal que se a aplicação $g\colon U \to \mathbb{R}^n$ é de classe C^1, com $|g-f|_1 \leq \delta$ em K, então $g|K$ é uma imersão.

10.5. Seja $f\colon U \to \mathbb{R}^n$ diferenciável. Dado um subconjunto $X \subset U$, diz-se que $f|X$ é um *mergulho* de X em $\mathbb{R}^n$ quando: 1º) $f|X$ é um homeomorfismo de X sobre $f(X)$; 2º) para cada $x \in X$, a derivada $f'(x)\colon \mathbb{R}^m \to \mathbb{R}^n$ é injetiva. Prove que se f é de classe C^1 e $K \subset U$ é um compacto convexo tal que $f|K$ é um mergulho então existe $\delta > 0$ tal que se $g\colon U \to \mathbb{R}^n$ de classe C^1, com $|g(x) - f(x)| < \delta$ e $|g'(x) - f'(x)| < \delta$ para todo $x \in K$ então $g|K$ é um mergulho.

§11.

11.1. Dê exemplos, em $\mathbb{R}^2$, de uma aplicação C^∞ aberta que não é uma submersão e de uma aplicação C^∞ injetiva que não é imersão.

11.2. Se $f\colon U \to \mathbb{R}^3$ é de classe C^1 e tem posto 3 em todos os pontos do aberto $U \subset \mathbb{R}^4$ então $|f(x)|$ não assume valor máximo para $x \in U$.

11.3. Se $f\colon U \to \mathbb{R}^m$ é uma submersão de classe C^k $(k \geq 1)$, definida no aberto $U \subset \mathbb{R}^{m+p}$, e $g\colon f(U) \to \mathbb{R}^n$ é tal que $g \circ f\colon U \to \mathbb{R}^n$ é de classe C^k, então g é de classe C^k.

11.4. Para cada $a = (a_0, \dots, a_m) \in \mathbb{C}^{m+1} = \mathbb{R}^{2m+2}$, indique com p_a o polinômio complexo $a_0 + a_1 z + \cdots + a_m z^m$. Se z_0 é uma raiz simples do polinômio p_a, mostre que existem bolas abertas $B = B(a; \varepsilon)$ em $\mathbb{C}^{m+1}$ e $D = D(z_0; \delta)$ em $\mathbb{C}$ tais que, para todo $b \in B$, o polinômio p_b tem uma única raiz $z = z(b)$ em D, a qual é simples, e a aplicação $B \to \mathbb{R}^2$, definida por $b \mapsto z(b)$, é de classe C^∞. ["Toda raiz simples de um polinômio é uma função C^∞ dos coeficientes desse polinômio".]

11.5. Seja $f\colon \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^n$ de classe C^k $(k \geq 1)$. Para cada $x \in \mathbb{R}^m$, indique com $f_x\colon \mathbb{R}^n \to \mathbb{R}^n$ a aplicação dada por $f_x(y) = f(x, y)$. Suponha que existam $a \in \mathbb{R}^m$ e $b \in \mathbb{R}^n$ tais que $f(a, b) = b$ e a transformação linear $f'_a(b)\colon \mathbb{R}^n \to \mathbb{R}^n$ não tenha valor próprio igual a 1. Prove que existem abertos $U \ni a$ em $\mathbb{R}^m$ e $V \ni b$ em $\mathbb{R}^n$ e uma aplicação $h\colon U \to V$ de classe C^k tal que, para todo $x \in U$, $h(x)$ é o único ponto fixo de f_x em V.

11.6. Seja $f: U \to \mathbb{R}$ de classe C^1 no aberto $U \subset \mathbb{R}^m$. Se $0 \in \mathbb{R}$ é valor regular de f então, para cada compacto $K \subset U$, existe $\varepsilon = \varepsilon_K > 0$ tal que 0 é valor regular de $g|K$ para toda função $g: U \to \mathbb{R}$ de classe C^1 com $|g(x) - f(x)| < \varepsilon$ e $|g'(x) - f'(x)| < \varepsilon$ qualquer que seja $x \in K$.

§12.

12.1. Seja $f: U \to \mathbb{R}^n$ de classe C^k $(k \geq 1)$ no aberto $U \subset \mathbb{R}^m$. Se, num ponto $a \in U$, a derivada $f'(a): \mathbb{R}^m \to \mathbb{R}^m$ tem posto p, então existe um mergulho $\varphi: V \to U$ de classe C^∞, definido num aberto $V \subset \mathbb{R}^p$, tal que $f \circ \varphi: V \to \mathbb{R}^n$ é um mergulho.

§13.

13.1. Seja $f: U \to U$ de classe C^k $(k \geq 1)$ no aberto conexo $U \subset \mathbb{R}^m$. Se $f \circ f = f$, ponha $M = f(U)$ e prove que f tem posto constante numa vizinhança de M. Conclua que M é uma superfície de classe C^k.

13.2. Seja $M \subset \mathbb{R}^n$ uma superfície de classe C^k $(k \geq 1)$ e dimensão m. Considere o conjunto $TM = \{(x, v) \in \mathbb{R}^n \times \mathbb{R}^n; x \in M, v \in T_xM\}$. Mostre que TM é uma superfície de dimensão $2m$ e classe C^{k-1} (conhecida como "o fibrado tangente de M"). Se N é outra superfície e $f: M \to N$ é uma aplicação de classe C^k, defina a aplicação tangente $Tf: TM \to TN$, pondo $(Tf)(x, v) = (f(x), f'(x) \cdot v)$ e mostre que Tf é de classe C^{k-1}.

13.3. Seja TM o fibrado tangente de uma superfície M, de classe C^k $(k \geq 2)$. Mostre que a aplicação $\pi: TM \to M$, $\pi(x, v) = x$, é uma submersão de classe C^{k-1} e que todo ponto de M tem uma vizinhança V tal que $\pi^{-1}(V)$ é difeomorfo ao produto $V \times \mathbb{R}^m$ $(m = \dim .M)$. Mais precisamente, existe um difeomorfismo $\varphi_V: V \times \mathbb{R}^m \to \pi^{-1}(V)$ tal que $\pi \circ \varphi_V = \pi_1$, onde $\pi_1: V \times \mathbb{R}^m \to V$ é a projeção sobre a primeira coordenada.

13.4. Seja $M \subset \mathbb{R}^n$ uma superfície de classe C^k $(k \geq 2)$ e dimensão m. Considere o conjunto $TM^\perp = \{(x, v) \in \mathbb{R}^n \times \mathbb{R}^n; x \in M, v \in T_xM^\perp\}$. Prove que $TM^\perp$ é uma superfície de classe C^{k-1} e dimensão n (chamada o "fibrado normal"de M). Prove que a aplicação $\pi: TM^\perp \to M$, definida por $\pi(x, v) = x$, é uma submersão de classe C^{k-1} e que todo ponto de M possui uma vizinhança aberta V tal que $\pi^{-1}(V)$ é difeomorfo ao produto $V \times \mathbb{R}^{n-m}$.

13.5. Seja $M \subset \mathbb{R}^n$ uma superfície de classe C^k $(k \geq 2)$ e dimensão m. Considere o conjunto $T_1M = \{(x, v) \in \mathbb{R}^n \times \mathbb{R}^n; x \in M, v \in T_xM, |v| = 1\}$. Prove que T_1M é uma superfície de classe C^{k-1} e dimensão $2m - 1$ (conhecida como "fibrado tangente unitário"de M). Mostre que T_1M é compacta se, e somente se, M é compacta. Prove que a aplicação $\pi: T_1M \to M$, $\pi(x, v) = x$, é uma submersão de classe C^{k-1} e que todo ponto $x \in M$ possui uma vizinhança V tal que $\pi^{-1}(V)$ é difeomorfo a $V \times S^{m-1}$.

13.6. Seja $M \subset \mathbb{R}^n$ uma superfície de classe C^k $(k \geq 2)$ e dimensão m. Prove que o conjunto $\nu_1(M) = \{(x, v) \in \mathbb{R}^n \times \mathbb{R}^n; x \in M, v \in T_xM^\perp, |v| = 1\}$ é uma superfície de classe C^{k-1} e dimensão $n - 1$. Mostre que $\nu_1(M)$ é compacta se, e somente se, M é compacta. [$\nu_1(M)$ é conhecida como o "fibrado normal

unitário de M".] Prove que a aplicação $\pi\colon \nu_1(M) \to M$, $\pi(x,v) = x$, é uma submersão de classe C^{k-1} e que cada ponto de M possui uma vizinhança V tal que $\pi^{-1}(V)$ é difeomorfo a $V \times S^{n-m-1}$.

13.7. Sejam $M \subset \mathbb{R}^n$ uma superfície de classe C^k $(k \geq 2)$ e $TM^\perp$ seu fibrado normal. Defina a aplicação $f\colon TM^\perp \to \mathbb{R}^n$ pondo $f(x,v) = x + v$. Prove que f é de classe C^{k-1}. Seja $M_0 = \{(x,v) \in TM^\perp; v = 0\}$. Mostre que f aplica M_0 difeomorficamente sobre M. Mais ainda: todo ponto $(x,0) \in M_0$ tem uma vizinhança aberta em $TM^\perp$ que é transformada difeomorficamente por f sobre uma vizinhança aberta de x em $\mathbb{R}^n$. Conclua que, se M é compacta então existe $\varepsilon > 0$ tal que o aberto $A_\varepsilon = \{(x,v) \in TM^\perp; |v| < \varepsilon\}$ é aplicado difeomorficamente por f sobre um aberto $V_0(M) \subset \mathbb{R}^n$, o qual se chama a *vizinhança tubular* de raio ε de M em $\mathbb{R}^n$. Conclua que existe uma aplicação $\pi\colon V_0(M) \to M$, de classe C^{k-1}, tal que $\pi(\pi(x)) = \pi(x)$ para todo $x \in M$. Para todo $y \in V_0(M)$, mostre que $y - x \in T_xM^\perp$ $(x = \pi(y))$, ou seja, que $x = \pi(y)$ é a interseção com M do único segmento de reta normal a M que contém y e está contido em $V_\varepsilon(M)$.

13.8. O conjunto $G_{k,m}$ das transformações lineares $P\colon \mathbb{R}^m \to \mathbb{R}^n$, auto-adjuntas, de posto k, com $P \cdot P = P$, é uma superfície de classe C^∞ e dimensão $k(m-k)$ em $\mathbb{R}^{m^2}$.

§14.

14.1. Sejam $M, N \subset \mathbb{R}^n$ respectivamente os conjuntos das matrizes 3×3 de postos 1 e 2. Mostre que M e N são superfícies C^∞, com $\dim .M = 5$, $\dim .N = 8$, ambas orientáveis.

14.2. Sejam M, N superfícies orientáveis, $f\colon M \to N$ uma aplicação de classe C^1 e $y \in N$ um valor regular de f. Prove que $f^{-1}(y)$ é uma superfície orientável.

14.3. Para toda superfície $M \subset \mathbb{R}^n$ de classe C^k $(k \geq 2)$ o fibrado normal $TM^\perp$ é sempre uma superfície orientável, o mesmo ocorrendo com o fibrado normal unitário $\nu_1(M)$. Mais precisamente, $TM^\perp$ e $\nu_1(M)$ possuem orientações naturais induzidas por $\mathbb{R}^n$.

14.4. Seja $M \subset \mathbb{R}^{m+1}$ uma hiperfície conexa de classe C^k $(k \geq 2)$. Mostre que o fibrado normal unitário de M é uma superfície conexa $\nu_1(M)$ se, e somente se, M é não-orientável. Caso M seja orientável, prove que $\nu_1(M)$ tem duas componentes conexas. A projeção natural $\pi\colon \nu_1(M) \to M$ é um difeomorfismo local. Quando M é orientável, π, restrita a cada componente conexa de $\nu_1(M)$, é um difeomorfismo sobre M.

14.5. Seja $M \subset \mathbb{R}^n$ uma superfície de classe C^k $(k \geq 2)$. Mostre que o fibrado tangente TM é sempre uma superfície orientável. Mais precisamente, TM possui uma orientação natural. Munindo TM e TN de suas orientações naturais, se $f\colon M \to N$ é um difeomorfismo de classe C^k, então a aplicação tangente $Tf\colon TM \to TN$ é um difeomorfismo (de classe C^{k-1}) que preserva a orientação.

14.6. Seja $f\colon M \to N$ um difeomorfismo local. Se N é orientável então M é orientável. Caso M seja conexa, orientável e f seja sobrejetivo, a seguinte

condição é necessária e suficiente para que N seja orientável: dados quaisquer $x, y \in M$ com $f(x) = f(y)$ o isomorfismo $f'(y)^{-1} \circ f'(x) \colon T_xM \to T_yM$ preserva a orientação desses dois espaços vetoriais.

14.7. Suponha que exista uma superfície P^m, de classe C^∞ e dimensão m, com a seguinte propriedade: existe um difeomorfismo local sobrejetivo $f \colon S^m \to P^m$, de classe C^∞, tal que $f(x) = f(y)$ se, e somente se, $y = \pm x$. Prove que P^m é orientável se, e somente se, m é ímpar. [Nota: P^m existe e chama-se "espaço projetivo real de dimensão m".]

14.8. Seja M uma superfície não-orientável. Prove que existem duas vizinhanças parametrizadas conexas, $U, V \subset M$ tais que $U \cap V$ tem duas componentes conexas, uma das quais a mudança de coordenadas tem determinante jacobiano positivo enquanto na outra esse determinante é negativo.

Capítulo VI

Integrais Múltiplas

1 A definição de integral

Um bloco *m-dimensional* é um produto cartesiano

$$A = \prod_{i=1}^{m} [a_i, b_i] = [a_1, b_1] \times \cdots \times [a_m, b_m] \subset \mathbb{R}^m$$

de m intervalos compactos $[a_i, b_i]$, cada um dos quais se chama uma *aresta* do bloco A. Quando todas as arestas têm o mesmo comprimento $b_i - a_i = a$, o bloco chama-se um *cubo m-dimensional*. Quando $m = 1$, A é um intervalo; para $m = 2$, o bloco reduz-se a um retângulo e o cubo a um quadrado. O ponto cujas coordenadas são $c_i = (a_i + b_i)/2$ chama-se o *centro* do bloco A.

Os *vértices* do bloco $A = \prod\limits_{i=1}^{m} [a_i, b_i]$ são os pontos $p = (c_1, \ldots, c_m)$ onde, para cada $i = 1, \ldots, m$, tem-se $c_i = a_i$ ou $c_i = b_i$. As *faces* do bloco A são os produtos cartesianos $F = L_1 \times \cdots \times L_m$ tais que $L_i = \{a_i\}$, $L_i = \{b_i\}$ ou $L_i = [a_i, b_i]$ para cada $i = 1, \ldots, m$. Diz-se que a face F tem *dimensão* k quando há precisamente k valores de i para os quais $L_i = [a_i, b_i]$. Em particular, cada vértice do bloco A é uma face de dimensão zero, enquanto o próprio A é uma face de dimensão m. Uma face $F \neq A$ chama-se *própria*.

O *volume m-dimensional* do bloco $A = \prod\limits_{i=1}^{m} [a_i, b_i]$ é, por definição, vol. $A = \prod\limits_{i=1}^{m} (b_i - a_i)$. Se A é um cubo m-dimensional cujas arestas têm comprimento a, tem-se vol. $A = a^m$.

O bloco m-dimensional $A = \Pi[a_i, b_i]$ é um subconjunto compacto de $\mathbb{R}^m$, cujo interior é o produto cartesiano int. $A = \Pi(a_i, b_i)$ dos intervalos abertos (a_i, b_i), o qual denominamos *bloco m-dimensional aberto.* Por definição, o volume do bloco aberto é o mesmo do bloco fechado correspondente.

Uma *partição* do bloco $A = \prod_{i=1}^{m} [a_i, b_i]$ é um conjunto finito do tipo $P = P_1 \times \cdots \times P_m$, onde cada P_i é uma partição do intervalo $[a_i, b_i]$. Os elementos $v \in P$ são chamados os *vértices* da partição P.

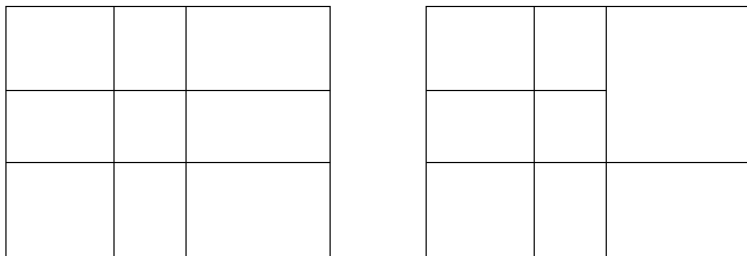
Uma partição $P = P_1 \times \cdots \times P_m$ do bloco A determina uma decomposição de A em sub-blocos do tipo $B = I_1 \times \cdots \times I_m$, onde cada I_j é um intervalo da partição P_j. Cada um desses sub-blocos B será chamado um *bloco da partição* P. Escreve-se então $B \in P$. Se para cada $j = 1, \ldots, m$, a partição P_j decompõe o intervalo $[a_j, b_j]$ em k_j subintervalos, então a partição P decompõe o bloco $A = \Pi[a_j, b_j]$ em $k_1 \cdot k_2 \cdot \cdots \cdot k_m$ sub-blocos.

Se B_1 e B_2 são blocos da mesma partição, então ou a interseção $B_1 \cap B_2$ é vazia ou é uma face k-dimensional comum a B_1 e B_2 $(k = 0, 1, \ldots, m-1)$.

Dada uma partição $P = P_1 \times \cdots \times P_m$ do bloco $A = \Pi[a_i, b_i]$, o comprimento de cada aresta $[a_i, b_i]$ é a soma dos comprimentos dos intervalos da partição P_i. Segue-se então da propriedade distributiva da multiplicação que o volume de A é a soma dos volumes de todos os blocos em que P decompõe A.

Sejam P e Q partições do bloco A. Diremos que Q é *mais fina* do que P quando $P \subset Q$. Se $P = P_1 \times \cdots \times P_m$ e $Q = Q_1 \times \cdots \times Q_m$, tem-se $P \subset Q$ se, e somente se, $P_1 \subset Q_1, \ldots, P_m \subset Q_m$. Neste caso, cada bloco da partição Q está contido num único bloco de P e cada bloco de P é a reunião dos blocos de Q neles contidos. Mais precisamente, se $P \subset Q$ então Q determina uma partição em cada bloco de P, logo o volume de um bloco $B \in P$ é a soma dos volumes dos blocos de Q nele contidos.

Na figura a seguir, vemos dois blocos (retângulos) decompostos como reunião de sub-blocos. A decomposição à esquerda provém de uma partição mas a da direita não.

Se $P = \Pi P_i$ e $Q = \Pi Q_i$ são partições do bloco A, a reunião $P \cup Q$ não é, em geral, uma partição de A. Mas existe uma partição $P + Q = \prod_{i=1}^{m} (P_i \cup Q_i)$ que refina P e Q.

A *norma* $|P|$ de uma partição $P = \Pi P_i$ é o maior comprimento de um subintervalo de qualquer das partições P_i, ou seja, é o maior comprimento das arestas dos blocos $B \in P$. Se tomarmos em $\mathbb{R}^m$ a norma do máximo, o diâmetro de um bloco será o comprimento da sua maior aresta. Neste caso, a norma da partição P será o maior diâmetro dos blocos $B \in P$.

Seja agora $f\colon A \to \mathbb{R}$ uma função real limitada, definida num bloco $A \subset \mathbb{R}^m$. Dada uma partição P de A, a cada bloco $B \in P$ associaremos os números

$$m_B = \text{inf.}\{f(x); x \in B\} \quad \text{e} \quad M_B = \text{sup.}\,\{f(x); x \in B\},$$

com os quais definiremos, respectivamente a *soma inferior* e a *soma superior* de f relativamente à partição P, pondo

$$s(f;P) = \sum_{B \in P} m_B \cdot \text{vol. } B \text{ e } S(f;P) = \sum_{B \in P} M_B \cdot \text{vol. } B,$$

os somatórios acima estendendo-se a todos os blocos B da partição P. Como $m_B \leq M_B$ para cada B, tem-se $s(f;P) \leq S(f;P)$.

Como no caso de funções de uma variável, mostraremos a seguir que, refinando-se uma partição, a soma inferior não diminui e a soma superior não aumenta.

Teorema 1. *Sejam P, Q partições do bloco $A \subset \mathbb{R}^m$, com $P \subset Q$, e $f\colon A \to \mathbb{R}$ uma função limitada. Tem-se*

$$s(f;P) \leq s(f;Q) \leq S(f;Q) \leq S(f;P).$$

Demonstração: Basta provar a primeira desigualdade, pois a terceira é análoga e a segunda já foi estabelecida. Designemos genericamente

por B os blocos da partição P e por B' os de Q. Se $B' \subset B$ então $m_B \leq m_{B'}$. Como vol. $B = \sum_{B' \subset B}$ vol. B', temos

$$\begin{aligned} s(f;P) &= \sum_{B \in P} m_B \cdot \text{vol. } B = \sum_{B \in P} m_B \cdot \left[\sum_{B' \subset B} \text{vol. } B'\right] \\ &= \sum_{B \in P} \left[\sum_{B' \subset B} m_B \cdot \text{vol. } B'\right] \leq \sum_{B \in P} \left[\sum_{B' \subset B} m_{B'} \cdot \text{vol. } B'\right] \\ &= \sum_{B' \in Q} m_{B'} \cdot \text{vol. } B' = s(f;Q). \end{aligned}$$

Corolário. *Seja $f: A \to \mathbb{R}$ limitada. Para quaisquer partições P e Q do bloco A, tem-se $s(f;P) \leq S(f;Q)$.*

Com efeito, sendo $P \subset P+Q$ e $Q \subset P+Q$, o teorema acima dá: $s(f;P) \leq s(f;P+Q) \leq S(f;P+Q) \leq S(f;Q)$.

Qualquer partição P do bloco A refina a partição trivial, cujo único sub-bloco é A. Segue-se então do Teorema 1 que, se $m \leq f(x) \leq M$ para todo $x \in A$ então

$$m \cdot \text{vol. } A \leq s(f;P) \leq S(f;P) \leq M \cdot \text{vol. } A.$$

Definiremos a *integral inferior* $\underline{\int}_A f(x)\,dx$ e a *integral superior* $\overline{\int}_A f(x)\,dx$ da função $f: A \to \mathbb{R}$, limitada no bloco A, pondo

$$\underline{\int}_A f(x)\,dx = \sup_P .\, s(f;P) \quad \text{e} \quad \overline{\int}_A f(x)\,dx = \inf_P .\, S(f;P),$$

o supremo e o ínfimo acima estendendo-se a todas as partições P do bloco A. Resulta do corolário e da observação acima, que se $m \leq f(x) \leq M$ para todo $x \in A$ então

$$m \cdot \text{vol. } A \leq \underline{\int}_A f(x)\,dx \leq \overline{\int}_A f(x)\,dx \leq M \cdot \text{vol. } A.$$

Observação: Seja P_0 uma partição arbitrária do bloco A. Para obtermos as integrais inferior e superior de uma função limitada $f: A \to \mathbb{R}$, basta considerarmos as partições que refinam P_0. Ou seja, tem-se:

$$\underline{\int}_A f(x)\,dx = \sup_{P \supset P_0} .\, s(f;P) \quad \text{e} \quad \overline{\int}_A f(x)\,dx = \inf_{P \supset P_0} .\, S(f;P).$$

Com efeito, para toda partição Q do bloco A, existe uma partição P que refina Q e P_0, a saber, $P = Q + P_0$. Então $s(f;Q) \leq s(f;P)$ e $S(f;P) \leq S(f;Q)$, com $P \supset P_0$ e daí resulta a observação.

Seja $f\colon A \to \mathbb{R}$ limitada no bloco $A \subset \mathbb{R}^m$. Diremos que f é *integrável* quando sua integral inferior for igual à sua integral superior. Definiremos então a *integral* de f como

$$\int_A f(x)\,dx = \underline{\int}_A f(x)\,dx = \overline{\int}_A f(x)\,dx.$$

Esta definição tem como caso particular a integral de uma função limitada num intervalo $[a,b] \subset \mathbb{R}$. Como no caso de uma variável, tem-se o seguinte critério de integrabilidade.

Teorema 2. *A fim de que uma função limitada $f\colon A \to \mathbb{R}$ seja integrável no bloco $A \subset \mathbb{R}^m$ é necessário e suficiente que, para todo $\varepsilon > 0$ dado, se possa obter uma partição P de A tal que*

$$S(f;P) - s(f;P) < \varepsilon.$$

Demonstração: Sejam σ o conjunto das somas inferiores e Σ o conjunto das somas superiores de f. Sabemos que se $s \in \sigma$ e $S \in \Sigma$ então $s \leq S$, o que nos dá $\sup.\sigma \leq \inf.\Sigma$. Para que se tenha $\sup.\sigma = \inf.\Sigma$ (isto é, para que f seja integrável) é suficiente que, para todo $\varepsilon > 0$ dado, se possam encontrar $s = s(f;P) \in \sigma$ e $S = S(f;Q) \in \Sigma$ tais que $S(f;Q) - s(f;P) < \varepsilon$. Logo a condição enunciada no teorema é suficiente para a integrabilidade de f. Reciprocamente, se f é integrável, dado $\varepsilon > 0$, obtemos partições P' e P'' tais que $S(f;P'') - s(f;P') < \varepsilon$. Em seguida, pondo $P = P' + P''$, obtemos $S(f;P) - s(f;P) \leq S(f;P'') - s(f;P') < \varepsilon$. Portanto a condição é necessária.

Seja $f\colon X \to \mathbb{R}$ limitada no conjunto $X \subset \mathbb{R}^m$. Chamaremos de *oscilação de f no conjunto X* ao número

$$\omega_X = \omega(f;X) = \sup.\{|f(x) - f(y)|;\ x, y \in X\}.$$

Como no caso de funções de uma variável, se indicarmos respectivamente com m_X e M_X o ínfimo e o supremo dos valores $f(x)$ para $x \in X$, teremos $\omega_X = M_X - m_X$. Evidentemente, $X \subset Y \Rightarrow \omega(f;X) \leq \omega(f;Y)$.

Assim, para que a função limitada $f\colon A \to \mathbb{R}$ seja integrável, é necessário e suficiente que a todo $\varepsilon > 0$ dado se possa fazer corresponder

uma partição P do bloco A tal que

$$\sum_{B \in P} \omega_B \cdot \text{vol. } B < \varepsilon.$$

Corolário do Teorema 2. *Toda função contínua $f\colon A \to \mathbb{R}$ é integrável.*

Com efeito, sendo o bloco A compacto, f é uniformemente contínua. Consideremos em $\mathbb{R}^m$ a norma do máximo. Dado $\varepsilon > 0$, existe $\delta > 0$ tal que $x, y \in A$, $|x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon / \text{vol. } A$. Assim, se P é uma partição de A com $|P| < \delta$, temos $\omega_B < \varepsilon/\text{vol. } A$ para todo bloco $B \in P$. Portanto

$$\sum_{B \in P} \omega_B \cdot \text{vol. } B < \frac{\varepsilon}{\text{vol. } A} \sum_{B \in P} \text{vol. } B = \varepsilon.$$

Pelo Teorema 2, f é integrável.

Exemplo 1. Seja $X \subset A$ um subconjunto tal que X e $A - X$ são ambos densos no bloco A. (Isto é, todo subconjunto aberto de A contém pontos de X e de $A - X$. Um exemplo seria X = conjunto dos pontos de A cujas coordenadas são números racionais.) A função característica do subconjunto X é $\chi_X\colon A \to \mathbb{R}$, definida por $\chi_X(x) = 0$ se $x \notin X$ e $\chi_X(x) = 1$ se $x \in X$. Para toda partição P do bloco A, e para todo bloco $B \in P$, temos $m_B = 0$ e $M_B = 1$. Segue-se que $s(\chi_X; P) = 0$ e $S(\chi_X; P) = \text{vol. } A$, seja qual for a partição P. Assim, a função χ_X não é integrável.

Exemplo 2. Sejam A, B blocos em $\mathbb{R}^m$, com $B \subset \text{int.}\, A$ e $\chi_B\colon A \to \mathbb{R}$ a função característica de B. Afirmamos que χ_B é integrável e que $\int_A \chi_B(x)\, dx = \text{vol. } B$. Com efeito, seja P_0 uma partição de A que tem B como um dos seus blocos. Se $P \supset P_0$ então $s(\chi_B; P) = \text{vol. } B$. Por uma observação anterior, ao considerarmos as integrais inferior e superior de χ_B, podemos restringir-nos às partições P que refinam P_0. Segue-se então que $\underline{\int}_A \chi_B(x)\, dx = \text{vol. } B$. Para calcular a integral superior, sejam B' um bloco com $B \subset \text{int}\, B' \subset A$ e Q_0 uma partição de A que tem B' como um dos seus blocos. Se $P \supset Q_0$ então $S(\chi_B; P) \leq \text{vol. } B'$. Daí resulta que $\text{vol. } B \leq \overline{\int}_A \chi_B(x)\, dx \leq \text{vol. } B'$. Como vol. B' pode ser tomado arbitrariamente próximo de vol. B, segue-se que $\overline{\int}_A \chi_B(x)\, dx = \text{vol. } B$.

Teorema 3. *Sejam $f, g\colon A \to \mathbb{R}$ funções integráveis no bloco $A \subset \mathbb{R}^m$. Então*

a) *A função $f+g$ é integrável e $\int_A [f(x)+g(x)]dx = \int_A f(x)dx + \int_A g(x)dx$;*

b) *Para todo $c \in \mathbb{R}$, a função $c \cdot f$ é integrável e $\int_A c \cdot f(x)dx = c \cdot \int_A f(x)dx$;*

c) *Se $f(x) \geq 0$ para todo $x \in A$ então $\int_A f(x)dx \geq 0$. Equivalentemente: se $f(x) \leq g(x)$ para todo $x \in A$ então $\int_A f(x)dx \leq \int_A g(x)dx$;*

d) *A função $|f(x)|$ é integrável e $|\int_A f(x)dx| \leq \int_A |f(x)|dx$. Em particular, se $|f(x)| \leq K$ para todo $x \in A$ então $|\int_A f(x)dx| \leq K \cdot \text{vol. } A$;*

e) *Se f é contínua, existe $c \in A$ tal que $\int_A f(x)dx = f(c) \cdot \text{vol. } A$.*

Demonstração: a) Para todo bloco $B \subset A$, temos $m_B(f) + m_B(g) \leq m_B(f+g)$ e $M_B(f+g) \leq M_B(f) + M_B(g)$. Daí resulta que, para quaisquer partições P, Q do bloco A, vale

$$s(f;P) + s(g;P) \leq s(f+g;P) \leq S(f+g;Q) \leq S(f;Q) + S(g;Q).$$

Portanto

$$\int_A f(x)dx + \int_A g(x)dx \leq \underline{\int}_A [f(x)+g(x)]dx \leq$$
$$\leq \overline{\int}_A [f(x)+g(x)]dx \leq \int_A f(x)dx + \int_A g(x)dx.$$

Assim, $f+g$ é integrável e $\int_A [f(x)+g(x)]dx = \int_A f(x)dx + \int_A g(x)dx$.

b) Vale $s(c \cdot f; P) = c \cdot s(f;P)$ se $c \geq 0$, enquanto $s(c \cdot f;P) = c \cdot S(f;P)$ para $c < 0$. [Cfr. Lema 3, pág. 311 do vol. 1.] Assim, quando $c < 0$, temos $\overline{\int}_A (c \cdot f)(x)dx = \inf_P . S(c \cdot f; P) = \inf_P . c \cdot s(f;P) = c \cdot \sup_P . s(f;P) = c \int_A f(x)dx$. De modo análogo, veríamos que $\underline{\int}_A (c \cdot f)(x)dx = c \int_A f(x)dx$, logo $c \cdot f$ é integrável e $\int_A (c \cdot f)(x)dx = c \cdot \int_A f(x)dx$. O caso $c > 0$ se trata da mesma maneira.

c) Se $f(x) \geq 0$ para todo $x \in A$ então $m_B \geq 0$ para todo bloco $B \subset A$, donde $s(f;P) \geq 0$ para toda partição P. Conseqüentemente,

$\underline{\int}_A f(x)dx = \sup_P . s(f;P) \geq 0$. Se $f(x) \leq g(x)$ para todo $x \in A$ então $(g - f)(x) \geq 0$ para todo x, logo $\int_A g(x)dx - \int_A f(x)dx = \int_A (g - f)(x)dx \geq 0$, ou seja, $\int_A f(x)dx \leq \int_A g(x)dx$.

d) De $||f(x)| - |f(y)|| \leq |f(x) - f(y)|$ resulta que a oscilação de $|f|$ em qualquer conjunto é menor do que ou igual à oscilação de f, logo f integrável implica $|f|$ integrável pelo Teorema 2. Além disso, como $-|f(x)| \leq f(x) \leq |f(x)|$ para todo $x \in A$, resulta do item anterior que

$$-\int_A |f(x)|dx \leq \int_A f(x)dx \leq \int_A |f(x)|dx,$$

o que significa $|\int_A f(x)dx| \leq \int_A |f(x)|dx$. Se for $|f(x)| \leq K$ para todo $x \in A$ então $|\int_A f(x)dx| \leq \int_A |f(x)|dx \leq \int_A K \cdot dx = K \cdot \text{vol. } A$.

e) Sejam $m = \inf.\{f(x); x \in A\}$ e $M = \sup.\{f(x); x \in A\}$. Sabemos que $m \leq \int_A f(x)dx/\text{vol. } A \leq M$. Logo, pelo Teorema do Valor Intermediário (A é conexo!), existe $c \in A$ tal que

$$\frac{\int_A f(x)dx}{\text{vol. } A} = f(c).$$

Os dois primeiros itens do teorema acima dizem que o conjunto das funções integráveis sobre um bloco $A \subset \mathbb{R}^m$ constitui um espaço vetorial real e a correspondência $f \mapsto \int_A f(x)dx$ é um funcional linear nesse espaço. O terceiro item diz que tal funcional é positivo e o quarto implica que o funcional $\int_A$ é contínuo quando se considera no espaço das funções integráveis a convergência uniforme.

O item c) é ainda complementado pela observação de que se $f \geq 0$ então $\int_A f(x)dx = 0$ só pode ocorrer quando $f(x) = 0$ para todo ponto $x \in A$ no qual a função f é contínua.

Toda função integrável $f\colon A \to \mathbb{R}$ se escreve como diferença $f = f_+ - f_-$ entre duas funções integráveis não-negativas. A função $f_+\colon A \to \mathbb{R}$ é chamada a *parte positiva* de f, enquanto que a função $f_-\colon A \to \mathbb{R}$ é sua *parte negativa*. Para todo $x \in A$, põe-se $f_+(x) = \max.\{f(x), 0\}$ e $f_-(x) = -\min.\{f(x), 0\}$. Assim, quando $f(x) \geq 0$, tem-se $f_+(x) = f(x)$ e $f_-(x) = 0$. Por outro lado, se $f(x) \leq 0$ então $f_+(x) = 0$ e $f_-(x) = -f(x)$. A igualdade $f = f_+ - f_-$ é evidente. Daí resulta que se f_+ e f_- são integráveis, f também é. Reciprocamente, como $f_+(x) = [f(x) + |f(x)|]/2$ e $f_-(x) = [|f(x)| - f(x)]/2$ para todo $x \in A$, a integrabilidade de f acarreta as de f_+ e f_-, pelo item d) do Teorema 3.

2 Conjuntos de medida nula

Como preparação para o teorema de Lebesgue (a ser demonstrado no §3), que dá uma condição necessária e suficiente para a integrabilidade de uma função limitada num bloco $A \subset \mathbb{R}^m$, estudaremos agora os conjuntos de medida nula no espaço euclidiano. Esses conjuntos possuem duas propriedades que justificam o interesse especial que por eles demonstramos aqui: são invariantes por aplicações de classe C^1 e têm interior vazio. A primeira propriedade permite definir conjuntos de medida nula em superfícies; a segunda confere aos conjuntos de medida nula um caráter de pequenez e faz com que seus complementares sejam não-vazios, fato importante para provar certas existências, principalmente quando coadjuvado pelo Teorema de Sard, que provaremos logo adiante.

Diremos que um conjunto $X \subset \mathbb{R}^m$ tem *medida nula*, e escreveremos med. $X = 0$, quando, para todo $\varepsilon > 0$ dado, for possível obter uma seqüência de cubos m-dimensionais abertos $C_1, C_2, \ldots, C_i, \ldots$ tais que $X \subset \bigcup_{i=1}^{\infty} C_i$ e $\sum_{i=1}^{\infty} \text{vol. } C_i < \varepsilon$.

Quando for necessário sermos mais precisos, diremos, nestas circunstâncias, que X tem *medida m-dimensional nula*, ou que X tem *medida nula em* $\mathbb{R}^m$.

Apresentaremos a seguir uma lista de dez proposições sobre os conjuntos de medida nula:

A) *Toda subconjunto de um conjunto de medida nula tem também medida nula.* (Evidente.)

B) *Toda reunião enumerável de conjuntos de medida nula é ainda um conjunto de medida nula.*

Demonstração: Seja $X = \bigcup_{i=1}^{\infty} X_i$, com med. $X_i = 0$ para todo $i \in \mathbb{N}$. Dado $\varepsilon > 0$ podemos obter, para cada $i \in \mathbb{N}$, uma seqüência de cubos abertos $C_{i1}, C_{i2}, \ldots, C_{ij}, \ldots$ tais que $X_i \subset \bigcup_{j=1}^{\infty} C_{ij}$ e $\sum_j \text{vol. } C_{ij} < \varepsilon/2^i$. Então X está contido na reunião (enumerável) de todos os C_{ij}. Dado qualquer subconjunto finito $F \subset \mathbb{N} \times \mathbb{N}$ existe $k \in \mathbb{N}$ tal que $(i,j) \in F \Rightarrow i \le k$, $j \le k$, logo

$$\sum_{(i,j)\in F} \text{vol.}\,(C_{ij}) \le \sum_{j=1}^{k} \left[\sum_{j=1}^{k} \text{vol.}\,(C_{ij}) \right] < \sum_{i=1}^{k} \varepsilon/2^i < \varepsilon.$$

Portanto, seja qual for a maneira de enumerar os C_{ij} numa seqüência, teremos $\sum\limits_{i,j} \text{vol.}\,(C_{ij}) \leq \varepsilon$. Assim, med. $X = 0$.

Em particular, como todo ponto tem evidentemente medida nula, todo conjunto enumerável é de medida nula.

C) *Seja $A \subset \mathbb{R}^m$ um bloco. Dada qualquer cobertura enumerável $A \subset \cup B_i$ por blocos abertos B_i, tem-se $\Sigma\,\text{vol.}\ B_i \geq \text{vol.}\ A$. Em particular, A não tem medida nula.*

Demonstração: Suponhamos inicialmente A fechado. Pelo Teorema de Borel-Lebesgue, existe k tal que $A \subset B_1 \cup \cdots \cup B_k$. Considerando um bloco fechado B, contendo $B_1, \ldots, B_k$, no qual estão definidas as funções características, temos $\chi_A \leq \chi_{B_1 \cup \cdots \cup B_k} \leq \chi_{B_1} + \cdots + \chi_{B_k}$. Segue-se do Exemplo 2 e do Teorema 3 (item c) que

$$\text{vol.}\ A = \int_B \chi_A(x)dx \leq \int_B \chi_{B_1}(x)dx + \cdots + \int_B \chi_{B_k}(x)dx =$$

$$= \text{vol.}\ B_1 + \cdots + \text{vol.}\ B_k \leq \sum_{i=1}^{\infty} \text{vol.}\ B_i\,.$$

Se o bloco A é aberto, temos vol. $D \leq \Sigma\,\text{vol.}\ B_i$ para todo bloco fechado $D \subset A$. Como vol. $A = \sup.\{\text{vol.}\ D; D \text{ bloco fechado contido em } A\}$, segue-se vol. $A \leq \Sigma\,\text{vol.}\ B_i$, o que completa a demonstração.

Decorre de A) e C) que todo conjunto de medida nula tem interior vazio pois se fosse int. $X \neq \varnothing$ então X conteria algum bloco, logo X não teria medida nula. A recíproca é falsa, conforme veremos agora.

Exemplo 3. (Conjunto de Cantor com medida positiva.) Exibiremos aqui um conjunto compacto $X \subset [0,1]$, com int. $X = \varnothing$ e med. $X \neq 0$. A construção de X é uma modificação do processo que leva ao conjunto de Cantor. (Veja págs. 172-178 e 339 do vol. 1.) Tomamos um número $a \in (0, 1/2)$, logo $\sum\limits_{n=1}^{\infty} a^n = 1 - \delta$, $\delta > 0$. Na primeira etapa da construção de X retiramos de $[0,1]$ um intervalo aberto de centro $1/2$ e comprimento a. Na segunda etapa, retiramos do centro de cada um dos dois intervalos fechados restantes um intervalo aberto de comprimento $a^2/2$, ficando quatro intervalos fechados. Depois de efetuada a n-ésima etapa da construção de X, restam 2^n intervalos fechados, dois a dois

disjuntos, de mesmo comprimento. A $(n+1)$-ésima etapa consiste em retirar do centro de cada um deles um intervalo aberto de comprimento $a^{n+1}/2^n$. A soma dos comprimentos dos intervalos retirados de $[0,1]$ na n-ésima etapa é igual a a^n. O conjunto X é o que resta de $[0,1]$ depois de efetuadas todas essas operações, ou seja $X = [0,1] - \cup J_s$ onde, indicando com $|J_s|$ o comprimento do intervalo J_s, temos $\sum_{s=1}^{\infty} |J_s| = \sum_{n=1}^{\infty} a^n = 1 - \delta$. Dada qualquer cobertura enumerável $X \subset \cup I_r$, por intervalos abertos I_r, temos $[0,1] \subset (\cup I_r) \cup (\cup J_s)$ logo, pelo item C) acima, $\Sigma|I_r| + \Sigma|J_s| \geq 1$, donde $\Sigma|I_r| \geq \delta$. Portanto X não tem medida nula. Notemos ainda que depois de cada etapa da construção de X, cada intervalo que resta tem comprimento menor do que a metade dos comprimentos dos intervalos que restaram da etapa anterior. Daí resulta que X não contém intervalos, isto é, int. $X = \varnothing$.

D) *Na definição de conjunto de medida nula, podemos usar cubos fechados, em vez de abertos.*

Demonstração: Seja $X \subset \cup C_i$, onde cada C_i é um cubo fechado, com $\Sigma\,\text{vol. } C_i < \varepsilon$. Ora, se $C = \prod_{j=1}^{m} [a_j, a_j + \ell]$ é um cubo fechado de aresta ℓ então, para cada $\delta > 0$,

$$D(\delta) = \prod_{j=1}^{m} \left(a_j - \frac{\delta}{2}, a_j + \ell + \frac{\delta}{2} \right)$$

é um cubo aberto de aresta $\ell + \delta$, contendo C. Evidentemente, $\lim_{\delta \to 0} \text{vol. } D(\delta) = \text{vol. } C$. Logo existe $\delta > 0$ tal que vol. $D(\delta) < 2\,\text{vol. } C$. Assim podemos obter, para cada i, um cubo aberto $D_i \supset C_i$, com vol. $D_i < 2 \cdot \text{vol. } C_i$, donde $X \subset \cup D_i$ e $\Sigma\,\text{vol. } D_i < 2 \cdot \Sigma\,\text{vol. } C_i = 2\varepsilon$. Isto prova a proposição.

E) *Se $X \subset \mathbb{R}^m$ é tal que, para todo $\varepsilon > 0$, existe uma seqüência de blocos fechados $A_1, \ldots, A_i, \ldots$ com $X \subset \cup A_i$ e $\Sigma\text{vol.} A_i < \varepsilon$, então* med. $X = 0$. *Ou seja: na definição de conjunto de medida nula podemos usar blocos fechados em vez de cubos.*

Demonstração: Mostremos inicialmente o seguinte: dados um bloco fechado $A \subset \mathbb{R}^m$ e $\varepsilon > 0$, existem cubos fechados $C_1, \ldots, C_k$ tais que

$A \subset \bigcup_{i=1}^{k} C_i$ e $\Sigma \, \text{vol.}\, C_i < \text{vol.}\, A + \varepsilon$. Com efeito, seja $A = \prod_{j=1}^{m} [a_j, b_j]$. Tomemos $q \in \mathbb{N}$. Para todo $j = 1, \ldots, m$ existe um inteiro $p_j \geq 0$ tal que $\frac{p_j}{q} < b_j - a_j \leq \frac{p_j + 1}{q} \cdot$ O bloco

$$A' = \prod_{j=1}^{m} \left[a_j, a_j + \frac{p_j + 1}{q} \right]$$

admite uma partição natural $P = P_1 \times \cdots \times P_m$, onde

$$P_j = \left\{ a_j, a_j + \frac{1}{q}, \ldots, a_j + \frac{p_j + 1}{q} \right\}.$$

Os blocos de P são cubos C_i, todos com arestas medindo $1/q$. Temos $A \subset A' = \cup C_i$, com

$$\Sigma \, \text{vol.}\, C_i = \text{vol.}\, A' = \prod_{j=1}^{m} \left(\frac{p_j}{q} + \frac{1}{q} \right) < \prod_{j=1}^{m} \left(b_j - a_j + \frac{1}{q} \right).$$

Como $\lim_{q \to \infty} \Pi \left(b_j - a_j + \frac{1}{q} \right) = \Pi(b_j - a_j) = \text{vol.}\, A$, podemos tomar q tão grande que $\Sigma \, \text{vol.}\, C_i < \text{vol.}\, A + \varepsilon$. Voltando ao enunciado da proposição tomamos, para cada $i \in \mathbb{N}$, cubos fechados C_{ij} tais que $A_i \subset \bigcup_j C_{ij}$, com $\sum_j \text{vol.}\, C_{ij} < \text{vol.}\, A_i + \frac{\varepsilon}{2^i}$, donde $X \subset \bigcup_{i,j} C_{ij}$, com $\sum_{i,j} \text{vol.}\, C_{ij} < 2\varepsilon$. Pelo item anterior, temos med. $X = 0$.

Com maior razão, podemos usar blocos abertos para definir medida nula.

Resulta de E) que se $m < n$ então, para qualquer $a \in \mathbb{R}^{n-m}$, todo subconjunto $X \subset \mathbb{R}^m \times \{a\} \subset \mathbb{R}^n$ tem medida nula em $\mathbb{R}^n$. Como $\mathbb{R}^m \times \{a\}$ é reunião enumerável de cubos m-dimensionais, basta provar isto para um desses cubos, digamos $C \times \{a\}$. Ora, para todo $\varepsilon > 0$, $C \times \{a\}$ está contido no bloco n-dimensional

$$A = C \times \prod_{i=1}^{n-m} \left(a_i - \frac{\varepsilon}{2}, \, a_i + \frac{\varepsilon}{2} \right)$$

com vol. $A = \varepsilon^{n-m} \cdot \text{vol.}\, C$. Segue-se de E) que med. $(C \times \{a\}) = 0$. Em particular, as faces próprias de um bloco m-dimensional têm medida nula em $\mathbb{R}^m$.

Em certas ocasiões, para provar que um conjunto tem medida nula, em vez de cobrí-lo com uma seqüência de blocos cuja soma dos volumes pode ser tomada arbitrariamente pequena, pode ser conveniente deixar "descoberto" um subconjunto que já sabemos ter medida nula. Este é o conteúdo da proposição abaixo. Note que o conjunto Y pode variar com ε.

F) *Seja $X \subset \mathbb{R}^m$. Se, para todo $\varepsilon > 0$, existe uma seqüência de blocos A_i (abertos ou fechados) tais que $\Sigma\,\text{vol.}\ A_i < \varepsilon$ e $X \subset (\cup A_i) \cup Y$, onde* med. $Y = 0$, *então X tem medida nula.*

Demonstração: Dado arbitrariamente $\varepsilon > 0$, obtemos os blocos A_i e o conjunto Y de medida nula, tais que $\Sigma\,\text{vol.}\ A_i < \varepsilon/2$ e $X \subset (\cup A_i) \cup Y$. Em seguida, obtemos novos blocos B_j, com $Y \subset \cup B_j$ e $\Sigma\,\text{vol.}\ B_j < \varepsilon/2$. Segue-se que $X \subset (\cup A_i) \cup (\cup B_j)$ com $\Sigma\,\text{vol.}\ A_i + \Sigma\,\text{vol.}\ B_j < \varepsilon$, logo X tem medida nula.

Uma aplicação $f\colon X \to \mathbb{R}^n$, definida num conjunto $X \subset \mathbb{R}^m$, diz-se *localmente Lipschitziana* quando, para todo $x \in X$, existem $V_x \subset \mathbb{R}^m$ aberto contendo x e $k_x > 0$ tais que $y, z \in V_x \Rightarrow |f(y) - f(z)| \le k_x \cdot |y - z|$. Noutras palavras, existe uma cobertura aberta $X \subset \cup V_x$ tal que cada restrição $f|(V_x \cap X)$ é Lipschitziana.

A proposição seguinte mostra que a noção de conjunto de medida nula é invariante por aplicações localmente Lipschitzianas. Nela, é importante notar que X e $f(X)$ devem estar no mesmo espaço euclidiano $\mathbb{R}^m$.

G) *Se $X \subset \mathbb{R}^m$ tem medida nula e $f\colon X \to \mathbb{R}^m$ é localmente Lipschitziana então $f(X)$ tem medida nula em $\mathbb{R}^m$.*

Demonstração: Consideremos primeiro o caso em que f é Lipschitziana. Então $|f(x) - f(y)| \le k \cdot |x - y|$ com $k > 0$ constante e $x, y \in X$ quaisquer. Tomemos em $\mathbb{R}^m$ a norma do máximo. Dado $\varepsilon > 0$, existe uma cobertura enumerável $X \subset \cup C_i$, onde cada C_i é um cubo de aresta a_i e $\sum_{i=1}^{\infty} a_i^m < \varepsilon/k^m$. Para cada $i \in \mathbb{N}$, temos $x, y \in X \cap C_i \Rightarrow |x - y| < a_i \Rightarrow |f(x) - f(y)| < k \cdot a_i$. Segue-se que cada uma das m projeções de $f(X \cap C_i)$ sobre os eixos está contida num intervalo de comprimento $k \cdot a_i$. Logo $f(X \cap C_i)$ está contido no produto cartesiano desses intervalos, que é um cubo D_i, cujo volume é igual a $k^m \cdot a_i^m$. Portanto

$f(X) = \bigcup_i f(X \cap C_i) \subset \cup D_i$ com $\sum_i$ vol. $D_i = k^m \cdot \sum_i a_i^m < \varepsilon$. Portanto med. $f(X) = 0$. No caso geral, temos $X \subset \cup V_x$, onde cada V_x é aberto e a restrição $f|(V_x \cap X)$ é Lipschitziana. Pelo Teorema de Lindelöf, tomamos uma subcobertura enumerável $X \subset \bigcup_{j=1}^{\infty} V_j$. Pela primeira parte, $f(V_j \cap X)$ tem medida nula, para cada $j \in \mathbb{N}$. Logo $f(X) = \bigcup_j f(V_j \cap X)$ é uma reunião enumerável de conjuntos de medida nula, donde med. $f(X) = 0$.

H) *Seja* $f: U \to \mathbb{R}^m$ *de classe* C^1 *no aberto* $U \subset \mathbb{R}^m$. *Se* $X \subset U$ *tem medida nula em* $\mathbb{R}^m$ *então* $f(X) \subset \mathbb{R}^m$ *também tem medida nula.*

Demonstração: Para cada $x \in X$, sejam V_x uma bola de centro x, com $\overline{V}_x \subset U$ e $k_x = \sup.\{|f'(y)|; y \in \overline{V}_x\}$. Pela Desigualdade do Valor Médio, $|f(y) - f(z)| \leq k_x \cdot |y - z|$ para quaisquer $y, z \in V_x$, logo f é localmente Lipschitziana. Assim, H) segue-se de G).

I) *Se* $m < n$ *e* $f: U \to \mathbb{R}^n$ *é de classe* C^1 *no aberto* $U \subset \mathbb{R}^m$ *então* $f(U)$ *tem medida nula em* $\mathbb{R}^n$.

Demonstração: Consideremos $0 \in \mathbb{R}^{n-m}$. Pela observação que se segue a E), $U \times \{0\}$ tem medida nula em $\mathbb{R}^n$. No aberto $W = U \times \mathbb{R}^{n-m} \subset \mathbb{R}^n$ definamos a aplicação $g: W \to \mathbb{R}^n$, de classe C^1, pondo $g(x, y) = f(x)$. Em virtude de H), o conjunto $g(U \times \{0\}) = f(U)$ tem medida nula em $\mathbb{R}^n$.

A proposição I) mostra, em particular, que não existem "curvas de Peano"em classe C^1, isto é, o desagradável fenômeno de uma aplicação f, definida num subconjunto $X \subset \mathbb{R}^m$, ter contido em sua imagem um cubo de $\mathbb{R}^n$, com $m < n$, pode ocorrer em classe C^0 mas não em classe C^1. (V. também Exerc. 2.2, adiante.)

Diremos que um conjunto $X \subset \mathbb{R}^m$ é *localmente de medida nula* quando, para cada $x \in X$, existir V_x aberto em $\mathbb{R}^m$, contendo x, tal que med. $(V_x \cap X) = 0$. Da cobertura aberta $X \subset \cup V_x$ extraimos, pelo Teorema de Lindelöf, uma subcobertura enumerável $X \subset \cup V_i$, logo $X = \bigcup_i (V_i \cap X)$ é uma reunião enumerável de conjuntos de medida nula, donde med. $X = 0$. Assim, um conjunto $X \subset \mathbb{R}^m$ é localmente de medida nula se, e somente se, tem medida nula.

Exemplo 4. Seja $M \subset \mathbb{R}^n$ uma superfície de classe C^1 e dimensão $m < n$ em $\mathbb{R}^n$. Dada uma parametrização $\varphi\colon U_0 \to U$ em M, segue-se de I) que a vizinhança parametrizada $U \subset M$ tem medida nula em $\mathbb{R}^n$. Como $U = A \cap M$, A aberto em $\mathbb{R}^n$, vemos que M é localmente de medida nula e por conseguinte, med. $M = 0$ em $\mathbb{R}^n$. Mais geralmente, se $X \subset \mathbb{R}^n$ é uma reunião enumerável de superfícies de classe C^1 e dimensões $< n$ então med. $X = 0$ em $\mathbb{R}^n$.

Definiremos agora conjuntos de medida nula numa superfície M, de classe C^1 e dimensão m.

Se o subconjunto $X \subset M$ estiver contido na imagem de uma parametrização $\varphi\colon U_0 \to U$, diremos que X tem *medida nula em* M quando $\varphi^{-1}(X)$ tiver medida nula em $\mathbb{R}^m$, isto é, quando $X = \varphi(X_0)$, com med. $(X_0) = 0$ em $\mathbb{R}^m$. Se $\psi\colon V_0 \to V$ for outra parametrização de classe C^1 em M, com $X \subset V$ então $\psi^{-1}(X) = (\psi^{-1} \circ \varphi)(\varphi^{-1}(X))$ tem ainda medida nula em $\mathbb{R}^m$ em virtude de H), pois $\psi^{-1} \circ \varphi\colon \varphi^{-1}(U \cap V) \to \mathbb{R}^m$ é de classe C^1 num aberto de $\mathbb{R}^m$.

No caso geral, diremos que um conjunto $X \subset M$ tem *medida nula* em M quando, para toda parametrização $\varphi\colon U_0 \to U$ de classe C^1, $U \cap X$ tiver medida nula em M de acordo com a definição acima.

Para que $X \subset M$ tenha medida nula na superfície M, basta que exista uma cobertura $X \subset \cup V_i$ (que podemos sempre supor enumerável, pelo Teorema de Lindelöf) por imagens de parametrizações $\varphi_i\colon V_{i0} \to V_i$, tais que $\varphi_i^{-1}(V_i \cap X)$ tenha medida nula em $\mathbb{R}^m$, para cada i. Com efeito, se $\varphi\colon U_0 \to U$ é outra qualquer parametrização de classe C^1 em M, temos $U \cap X = \bigcup_i W_i$, onde $W_i = U \cap V_i \cap X$, logo

$$\varphi^{-1}(U \cap X) = \bigcup_i \varphi^{-1}(W_i) = \bigcup_i (\varphi^{-1} \circ \varphi_i)(\varphi_i^{-1}(W_i))$$

tem medida nula em $\mathbb{R}^m$, pois cada $\varphi_i^{-1}(W_i) \subset \varphi_i^{-1}(V_i \cap X)$ tem medida nula.

Os conjuntos de medida nula numa superfície gozam de propriedades análogas às de seus homônimos do espaço euclidiano, quais sejam:

a) Se med. $X = 0$ em M e $Y \subset X$ então med. $Y = 0$ em M;

b) Se $X_1, \ldots, X_i, \ldots$ têm medida nula em M então $X = \cup X_i$ tem medida nula em M;

c) Todo conjunto X de medida nula em M tem interior vazio em M, ou seja, $M - X$ é denso em M. Em particular, $M - X$ não é vazio;

d) Se $f: M \to N$ é de classe C^1, com dim. $M =$ dim. N e $X \subset M$ tem medida nula em M então $f(X)$ tem medida nula em N;

e) Se dim. $M <$ dim. N e $f: M \to N$ é de classe C^1 então $f(M)$ tem medida nula em N.

Em relação à propriedade c), convém observar que a superfície M pode ser reunião de dois conjuntos com interior vazio. Por exemplo, temos $\mathbb{R} = \mathbb{Q} \cup (\mathbb{R} - \mathbb{Q})$. Mas, em virtude de b), se X e Y têm medida nula em M então $X \cup Y$ tem medida nula e portanto interior vazio em M. Em particular, $(M - X) \cap (M - Y) \neq \varnothing$. Mais geralmente, dada uma infinidade enumerável de conjuntos com medida nula na superfície M, existem pontos de M que não pertencem a nenhum deles.

Nossa décima proposição sobre conjuntos de medida nula será o Teorema de Sard, um resultado "inevitável"sobre aplicações diferenciáveis. Ele será usado de modo essencial nos parágrafos 9 e 10 do capítulo seguinte. Em sua forma geral, o Teorema de Sard se refere a aplicações entre superfícies de dimensões m e n quaisquer. Consideraremos apenas o caso particular em que $m = n$, o que tornará a prova muito mais fácil e será suficiente para os nossos propósitos. Para apresentações mais completas do assunto, vejam-se [16] e [14].

Na demonstração abaixo, usaremos alguns resultados que, embora elementares e plausíveis, só serão provados mais adiante, neste mesmo capítulo.

Teorema de Sard. *Seja $f: M \to N$ uma aplicação de classe C^1 entre superfícies da mesma dimensão m. Seja S o conjunto dos pontos $x \in M$ nos quais a derivada $f'(x): T_xM \to T_{f(x)}M$ não é um isomorfismo. Então $f(S)$ tem medida nula em N.*

Os pontos do conjunto S chamam-se *pontos singulares de f*. Portanto o Teorema de Sard se enuncia, em palavras, assim: a imagem do conjunto dos pontos singulares por uma aplicação de classe C^1 tem medida nula.

Note que $N - f(S)$ é o conjunto dos valores regulares de f. Uma conseqüência importante do Teorema de Sard é que $f(S)$ tem interior vazio, donde $N - f(S)$ é denso em N. Em particular, $N - f(S) \neq \varnothing$, isto é, toda aplicação $f: M \to N$ de classe C^1 entre superfícies de mesma dimensão possui valores regulares. Na verdade, "quase todos"os pontos $y \in N$ são valores regulares de f.

Demonstração: Para cada $x \in S$, tomemos parametrizações $\psi\colon V_0 \to V$ em N, com $f(x) \in V$, e $\varphi\colon U_0 \to U$ em M, com $x \in U$ e $f(U) \subset V$. Basta provar que, para todo $x \in S$, $f(U \cap S)$ tem medida nula em N. [De fato, da cobertura $S \subset \cup U$ extraímos, pelo Teorema de Lindelöf, uma subcobertura enumerável $S \subset \cup U_i$ e

$$f(S) = f(\cup(U_i \cap S)) = \cup f(U_i \cap S)$$

será uma reunião enumerável de conjuntos de medida nula em N.] Por outro lado, med. $f(U \cap S) = 0$ em $N \Leftrightarrow$ med. $\psi^{-1} f(U \cap S) = 0$ em $\mathbb{R}^m \Leftrightarrow$ med. $[\psi^{-1} f \varphi(\varphi^{-1}(U \cap S))] = 0$ em $\mathbb{R}^m$. Como $\varphi^{-1}(U \cap S)$ é o conjunto dos pontos singulares da aplicação $\psi^{-1} \circ f \circ \varphi\colon U_0 \to \mathbb{R}^m$, o Teorema de Sard se reduz a provar o seguinte:

Seja $f\colon U \to \mathbb{R}^m$ uma aplicação de classe C^1 no aberto $U \subset \mathbb{R}^m$. Seja $S = \{x \in U; \text{det. } f'(x) = 0\}$. Então $f(S)$ tem medida nula em $\mathbb{R}^m$.

Novamente pelo Teorema de Lindelöf, U é reunião enumerável de cubos fechados. Logo, basta provar que se C é um cubo fechado, de aresta $a > 0$, contido em U, e $T = \{x \in C; \text{det. } f'(x) = 0\}$ então $f(T)$ tem medida nula em $\mathbb{R}^m$.

Tomemos no espaço $\mathbb{R}^m$ a norma euclidiana.

Subdividindo cada uma de suas arestas em k partes iguais, obtemos uma partição de C, cujos blocos são k^m pequenos cubos C_i, de mesma aresta $\dfrac{a}{k} = \delta$ e volume $\dfrac{a^m}{k^m} = \delta^m$. Se $x, y \in C_i$, temos $|x - y| \leq m \cdot \delta$.

Em cada pequeno cubo C_i tal que $C_i \cap T \neq \varnothing$, escolhamos um ponto $x_i \in C_i \cap T$. A imagem da transformação linear $f'(x_i)\colon \mathbb{R}^m \to \mathbb{R}^m$ está contida num subespaço vetorial $E_i \subset \mathbb{R}^m$, de dimensão $m-1$. Todos os pontos $f(x_i) + f'(x_i) \cdot v$, $v \in \mathbb{R}^m$, pertencem à variedade afim $L_i = f(x_i) + E_i$, de dimensão $m-1$ em $\mathbb{R}^m$. Para cada $x \in C_i$ podemos escrever

$$f(x) = f(x_i) + f'(x_i) \cdot (x - x_i) + r_i(x).$$

Dado arbitrariamente $\varepsilon > 0$, o inteiro k pode ser tomado tão grande que, para todo cubo C_i contendo pontos de T e todo $x \in C_i$ valha $|r_i(x)| \leq \varepsilon|x - x_i| \leq m \cdot \delta \cdot \varepsilon$. Ora, para todo $x \in C_i$, se $c = \sup.\{|f'(x)|; x \in C\}$, temos $|f'(x_i) \cdot (x - x_i)| \leq c|x - x_i| < m \cdot c \cdot \delta$, logo, para todo $x \in C_i$, o ponto $f(x_i) + f'(x_i) \cdot (x - x_i)$ pertence a um cubo de centro $f(x_i)$ e aresta $2m\delta c$ em L_i. Se considerarmos o paralelepípedo retangular P_i em $\mathbb{R}^m$ que tem esse cubo como seção média e altura $2m\varepsilon\delta$, vemos que

vol. $P_i = 2^m \cdot m^m \cdot c^{m-1} \cdot \delta^m \cdot \varepsilon = A \cdot \delta^m \cdot \varepsilon$. A imagem $f(T)$ está contida na reunião de, no máximo, k^m desses paralelepípedos P_i, cuja soma dos volumes não excede $A \cdot k^m \cdot \delta^m \cdot \varepsilon = A \cdot a^m \cdot \varepsilon$. Como $\varepsilon > 0$ é arbitrário, vemos que $f(T)$ tem medida nula.

Comentário sobre a demonstração acima. Até agora, os únicos conjuntos cujos volumes sabemos calcular são os blocos retangulares. Na realidade, nem definimos ainda o que é o "volume"de um conjunto. Isto será feito logo em seguida, de tal forma que o volume de um paralelepípedo retangular seja o produto do volume da base pelo comprimento da altura, justificando o argumento que provou o Teorema de Sard. Além disso, ficará claro a partir da definição que se $X \subset \mathbb{R}^m$ é tal que, para todo $\varepsilon > 0$, existe uma cobertura $X \subset \cup P_i$, onde a soma dos volumes dos conjuntos P_i é $< \varepsilon$, então X é um conjunto de medida nula. Isto também foi utilizado na demonstração precedente.

3 Caracterização das funções integráveis

Na demonstração do Teorema de Lebesgue, será utilizado o conceito de oscilação de uma função num ponto. Esta noção, que já foi apresentada no Vol. 1 (página 340) no caso particular de subconjuntos da reta e funções neles definidas, será retomada agora no contexto do espaço euclidiano $\mathbb{R}^m$.

Seja $f\colon X \to \mathbb{R}$ limitada no conjunto $X \subset \mathbb{R}^m$. Fixemos $x \in X$ e, para cada $\delta > 0$, ponhamos $\Omega(\delta) = \omega[f; X \cap B(x;\delta)] =$ oscilação de f no conjunto dos pontos de X que distam menos de δ do ponto x. Isto define uma função não-negativa $\Omega\colon (0, +\infty) \to \mathbb{R}$. Como f é limitada, Ω também é. Além disso, $\delta \leq \delta' \Rightarrow \Omega(\delta) \leq \Omega(\delta')$. Logo existe o limite

$$\omega(f;x) = \lim_{\delta \to 0} \omega[f; X \cap B(x;\delta)] = \inf_{\delta > 0} \omega[f; X \cap B(x;\delta)],$$

que chamaremos a *oscilação da função f no ponto x*. Quando não houver possibilidade de confusão, escreveremos simplesmente $\omega(x)$, em vez de $\omega(f;x)$.

A oscilação goza das seguintes propriedades:

I. $\omega(f;x) \geq 0$ *para todo* $x \in X$;

II. $\omega(f;x) = 0$ *se, e somente se, f é contínua no ponto x;*

III. *Se* $x \in \text{int}.Y$ *e* $Y \subset X$ *então* $\omega(f;x) \leq \omega(f;Y)$*;*

IV. *Se* $\omega(f;x) < c$ *então existe* $\delta > 0$ *tal que* $\omega(f;y) < c$ *para todo* $y \in X$ *com* $|y - x| \leq \delta$*;*

V. *Se* $X \subset \mathbb{R}^m$ *é fechado (respect. compacto) então, para todo* $c \geq 0$ *o conjunto* $\{x \in X; \omega(x) \geq c\}$ *é fechado (respect. compacto).*

Vamos às demonstrações:

I é evidente.

Quanto a II, a afirmação $\omega(f;x) = 0$ significa que, para todo $\varepsilon > 0$ dado, pode-se obter $\delta > 0$ tal que $y, z \in X \cap B(x;\delta) \Rightarrow |f(y) - f(z)| \leq \varepsilon$. Isto é claramente equivalente à continuidade de f no ponto x.

Para provar III, notemos que existe $\delta > 0$ tal que $B(x;\delta) \subset Y$. Logo

$$\omega(f;x) \leq \omega[f; X \cap B(x;\delta)] = \omega[f; B(x;\delta)] \leq \omega(f;Y).$$

No que diz respeito a IV, pela definição de limite, existe $\delta > 0$ tal que $\omega[f; X \cap B(x;\delta)] < c$. Dado $y \in X \cap B(x;\delta)$, tomamos $\eta > 0$ tal que $B(y;\eta) \subset B(x;\delta)$ e obtemos

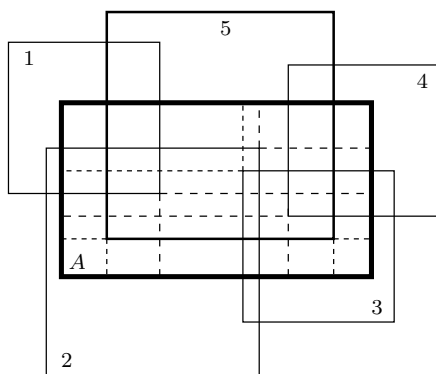
$$\omega(f;y) \leq \omega[f; X \cap B(y;\eta)] \leq \omega[f; X \cap B(x;\delta)] < c.$$

Finalmente, se X é fechado e uma seqüência de pontos $x_k \in X$ é tal que $\omega(f;x_k) \geq c$ para todo $k \in \mathbb{N}$ e $\lim x_k = x$ então, em primeiro lugar, $x \in X$. Em segundo lugar, não pode ser $\omega(f;x) < c$ por causa de IV. Logo $\omega(f;x) \geq c$ e o conjunto $\{x \in X; \omega(f;x) \geq c\}$ é, portanto, fechado em $\mathbb{R}^m$. Se X for compacto, esse conjunto será também limitado, donde compacto, o que prova V.

Estabelecidos estes preliminares, provemos o teorema de caracterização das funções integráveis.

Teorema 4 (Lebesgue). *Uma função* $f\colon A \to \mathbb{R}$*, limitada no bloco* $A \subset \mathbb{R}^m$*, é integrável se, e somente se, o conjunto* D_f *dos seus pontos de descontinuidade tem medida nula.*

Demonstração: Suponhamos primeiro que med. $D_f = 0$. Dado arbitrariamente $\varepsilon > 0$, tomemos uma cobertura enumerável $D_f \subset \cup C'_i$ por cubos abertos tais que $\Sigma \,\text{vol.}\, C'_i < \varepsilon/2K$, onde $K = M - m$ (diferença entre o sup. e o inf. de f em A) é a oscilação de f no bloco A. Para cada ponto $x \in A - D_f$ tomemos um cubo aberto C''_x, contendo x, tal

que a oscilação de f em $\overline{C''_x} \cap A$ seja inferior a $\varepsilon/(2\,\text{vol.}\ A)$. Como A é compacto, da cobertura aberta $A \subset (\cup C'_i) \cup (\cup C''_x)$ podemos retirar uma subcobertura finita $A \subset C'_1 \cup \cdots \cup C'_r \cup C''_1 \cup \cdots \cup C''_s$. Seja P uma partição de A tal que cada bloco (aberto) $B \in P$ esteja contido, ou num dos cubos C'_i, ou num dos cubos C''_j. A partição P pode ser obtida "prolongando-se as faces" dos cubos C'_i e C''_j. Mais precisamente, se $A = \prod_{k=1}^{m} [a_k, b_k]$ então $P = P_1 \times \cdots \times P_m$ onde, para cada $k = 1, \ldots, m$, P_k é o conjunto formado por a_k, b_k e

mais as k-ésimas coordenadas dos vértices dos cubos C'_i e C''_j. Indicaremos genericamente com a notação B' os blocos de P que estão contidos em algum cubo C'_i. Os demais blocos de P (necessariamente contidos em cubos C''_j) serão chamados B''. A soma dos volumes dos blocos B' é menor do que $\varepsilon/2K$ e em cada bloco B'' a oscilação de f não excede $\varepsilon/(2\,\text{vol.}\ A)$. Portanto, a partição P nos dá

$$\sum_{B \in P} \omega_B \cdot \text{vol.}\ B = \sum_{B'} \omega_{B'} \cdot \text{vol.}\ B' + \sum_{B''} \omega_{B''} \cdot \text{vol.}\ B'' \leq$$
$$\leq K \cdot \Sigma\,\text{vol.}\ B' + \frac{\varepsilon}{2\,\text{vol.}\ A} \Sigma\,\text{vol.}\ B'' <$$
$$< K \cdot \frac{\varepsilon}{2 \cdot K} + \frac{\varepsilon}{2\,\text{vol.}\ A} \cdot \text{vol.}\ A = \varepsilon.$$

Logo f é integrável. Reciprocamente, suponhamos f integrável. Temos $D_f = D_1 \cup D_2 \cup \cdots \cup D_i \cup \ldots$ onde $D_i = \{x \in A; \omega(f; x) \geq 1/i\}$. Para mostrar que D_f tem medida nula provaremos que, para cada $i \in \mathbb{N}$, med. $D_i = 0$. Seja então dado $\varepsilon > 0$. Como f é integrável, existe

uma partição P do bloco A, tal que $\sum_{B\in P} \omega_B \cdot \text{vol. } B < \varepsilon/i$. Indiquemos genericamente com B' os blocos da partição P que contêm algum ponto de D_i em seu interior. Para cada um desses blocos, temos $\omega_{B'} \geq 1/i$, em vista da propriedade III da oscilação $\omega(f;x)$. Logo

$$\frac{1}{i}\,\Sigma\,\text{vol. } B' \leq \Sigma\,\omega_{B'} \cdot \text{vol. } B' \leq \sum_{B\in P} \omega_B \cdot \text{vol. } B < \frac{\varepsilon}{i}\cdot$$

Multiplicando por i, vem: $\Sigma\,\text{vol. } B' < \varepsilon$. Ora, é claro que $D_i \subset (\cup\, B') \cup Y$, onde Y é a reunião das faces próprias dos blocos $B \in P$ nas quais existe algum ponto de D_i. Sabemos que Y tem medida nula. Pela proposição F) do parágrafo anterior, temos med. $D_i = 0$, o que completa a demonstração.

Definiremos agora o volume (segundo Jordan) de um subconjunto limitado do espaço euclidiano.

Diz-se que o conjunto limitado $X \subset \mathbb{R}^m$ é *J-mensurável* (mensurável segundo Jordan) quando, tomando-se um bloco $A \subset \mathbb{R}^m$ que contenha X, a função característica $\chi_X \colon A \to \mathbb{R}$ é integrável. Quando X é J-mensurável, seu *volume* é, por definição, a integral de sua função característica:

$$\text{vol. } X = \int_A \chi_X(x)\,dx.$$

Como conseqüência do Teorema de Lebesgue, provaremos abaixo uma importante caracterização dos conjuntos J-mensuráveis. Antes, recordemos que a fronteira ∂X de um conjunto $X \subset \mathbb{R}^m$ é o conjunto dos pontos $a \in \mathbb{R}^m$ tais que toda vizinhança de a contém pontos de X e pontos de $\mathbb{R}^m - X$. Tem-se $\mathbb{R}^m = (\text{int.}\, X) \cup \partial X \cup \text{int.}(\mathbb{R}^m - X)$, reunião disjunta.

Teorema 5. *Um conjunto limitado $X \subset \mathbb{R}^m$ é J-mensurável se, e somente se, sua fronteira ∂X tem medida nula.*

Demonstração: Dado o bloco $A \supset X$ em $\mathbb{R}^m$, seja D o conjunto dos pontos de descontinuidade da função característica $\chi_X \colon A \to \mathbb{R}$. Evidentemente, $D \subset \partial X$. Por outro lado, um ponto de ∂X que não seja uma descontinuidade de χ_X deve pertencer a ∂A. Assim, podemos escrever $\partial X = D \cup (\partial X \cap \partial A)$. [Se $X \subset \text{int.}\, A$ então $\partial X = D$.] Como ∂A (reunião das faces próprias do bloco A) tem medida nula, segue-se que med. $\partial X = 0 \Leftrightarrow$ med. $D = 0$.

Em particular, o fato de X ser J-mensurável (e, como se vê facilmente, o valor de vol. X) independe do bloco A tomado na definição.

Exemplo 5. Um bloco é um conjunto J-mensurável e seu volume já foi determinado no Exemplo 2. Uma bola (aberta ou fechada) é J-mensurável pois sua fronteira é uma esfera, que tem medida nula, pelo Exemplo 4. De um modo geral, se $X \subset \mathbb{R}^m$ é um conjunto limitado, cuja fronteira é uma reunião enumerável de superfícies de classe C^1 e dimensões $< m$ então X é J-mensurável. Um conjunto de Cantor de medida positiva (Exemplo 3) não é J-mensurável, pois sua fronteira é ele próprio, que não tem medida nula.

Mais geralmente, dado um conjunto limitado $X \subset \mathbb{R}^m$, podemos considerar um bloco $A \supset X$ em $\mathbb{R}^m$ e definir, respectivamente, o *volume interno* de X e o *volume externo* de X pondo

$$\text{vol. int.}\, X = \underline{\int}_A \chi_X(x)\, dx, \quad \text{vol. ext.}\, X = \overline{\int}_A \chi_X(x)\, dx.$$

Recordando as definições da integral inferior e da integral superior, vemos que o volume interno de X é o supremo, relativamente a todas as partições P do bloco A, dos números $s(\chi_X; P) =$ soma dos volumes dos blocos de P que estão contidos em X, enquanto que o volume externo de X é o ínfimo, relativamente a todas as partições P do bloco A, dos números $S(\chi_X; P) =$ soma dos volumes dos blocos de P que têm algum ponto em comum com X.

Em particular, um conjunto limitado $X \subset \mathbb{R}^m$ (é J-mensurável e) tem volume zero se, e somente se, tem volume externo zero, ou seja, para todo $\varepsilon > 0$ dado, pode-se achar um bloco $A \supset X$ e uma partição P de A tal que soma dos volumes dos blocos de P que intersectam X é menor do que ε. Isto equivale a dizer que, dado $\varepsilon > 0$, existem blocos $B_1, \ldots, B_k$ em $\mathbb{R}^m$ com $X \subset B_1 \cup \cdots \cup B_k$ e $\Sigma\, \text{vol.}\, B_i < \varepsilon$.

Segue-se do Teorema de Borel-Lebesgue que um conjunto compacto tem volume zero se, e somente se, tem medida nula. Por exemplo, a fronteira ∂X de um conjunto limitado $X \subset \mathbb{R}^m$ é sempre compacta. Logo, as seguintes afirmações são equivalentes:

a) X é J-mensurável;

b) ∂X tem medida nula;

c) ∂X tem volume zero.

Exemplo 6. Um conjunto (não-compacto) de medida nula nem sempre é J-mensurável: o conjunto dos pontos de um bloco A que têm coordenadas racionais é enumerável, logo tem medida nula mas não é J-mensurável, como se viu no Exemplo 1. Mas se $X \subset \mathbb{R}^m$ é J-mensurável e tem medida nula (ou, mais geralmente, tem interior vazio) então seu volume deve ser zero pois, como nenhum bloco está contido em X, temos $\text{vol. int.}\, X = 0$. Assim, se $X \subset \mathbb{R}^m$ é J-mensurável, tem-se $\text{vol.}\ X = 0$ se, e somente se, X tem interior vazio. (Compare Exemplo 3.)

Teorema 6. *Sejam X, Y subconjuntos J-mensuráveis do bloco $A \subset \mathbb{R}^m$. Então:*

a) $X \cup Y$, $X \cap Y$ *e* $A - X$ *são J-mensuráveis;*

b) $\text{vol.}\,(X \cup Y) + \text{vol.}\,(X \cap Y) = \text{vol.}\ X + \text{vol.}\ Y$.

Demonstração: A afirmação a) segue-se imediatamente das três inclusões seguintes, as quais, por outro lado, decorrem da definição de fronteira: $\partial(X \cup Y) \subset \partial X \cup \partial Y$, $\partial(X \cap Y) \subset \partial X \cup \partial Y$ e $\partial(A - X) \subset \partial A \cup \partial X$. A afirmação b) resulta, por sua vez, da igualdade óbvia $\chi_{X\cup Y} + \chi_{X\cap Y} = \chi_X + \chi_Y$.

Corolário. *Se X, Y são J-mensuráveis e* $\text{int.}(X \cap Y) = \varnothing$ *então* $\text{vol.}\,(X \cup Y) = \text{vol.}\ X + \text{vol.}\ Y$.

Como $\text{int.}(X \cap Y) = \text{int.}\, X \cap \text{int.}\, Y$, a hipótese do corolário acima significa que os conjuntos J-mensuráveis X e Y têm em comum, no máximo, pontos de suas fronteiras. Nestas condições, o volume de $X \cup Y$ é a soma de $\text{vol.}\ X$ e $\text{vol.}\ Y$.

Seja M uma superfície de classe C^1. Um conjunto $X \subset M$ diz-se *J-mensurável* em M quando sua fronteira em M (conjunto dos pontos $x \in M$ tais que $x = \lim x_k$, $x_k \in X$ e $x = \lim y_k$, $y_k \in M - X$) é um conjunto de medida nula em M. A reunião e a interseção de dois subconjuntos J-mensuráveis em M é ainda J-mensurável, o mesmo ocorrendo com o complemento $M - X$ de um conjunto J-mensurável X. O volume de um conjunto J-mensurável $X \subset M$ só será definido no Capítulo VII, quando estivermos de posse da noção de integral de superfície.

Definiremos agora a integral de uma função limitada $f\colon X \to \mathbb{R}$, cujo domínio é um conjunto J-mensurável $X \subset \mathbb{R}^m$.

Para definirmos $\int_X f(x)\,dx$, consideraremos um bloco $A \subset \mathbb{R}^m$ que contenha o conjunto X e estenderemos f a uma função $\overline{f}\colon A \to \mathbb{R}$, pondo $\overline{f}(x) = 0$ se $x \in A - X$ e $\overline{f}(x) = f(x)$ quando $x \in X$.

Diremos que $f\colon X \to \mathbb{R}$ é integrável quando a função $\overline{f}\colon A \to \mathbb{R}$ for integrável e poremos, por definição,

$$\int_X f(x)\,dx = \int_A \overline{f}(x)\,dx.$$

Observação: Às vezes se escreve $\overline{f} = f \cdot \chi_X$, o que significa $\overline{f}(x) = f(x) \cdot \chi_X(x)$ para todo $x \in A$. Esta igualdade é um abuso de notação, que corresponde a considerar um produto igual a zero quando um dos fatores for zero e o outro não estiver definido. O que é verdadeiro (e óbvio) é que $\overline{f} = \overline{f} \cdot \chi_X$.

Teorema 7. *Seja $X \subset \mathbb{R}^m$ um conjunto J-mensurável. Uma função limitada $f\colon X \to \mathbb{R}$ é integrável se, e somente se, o conjunto D_f dos seus pontos de descontinuidade tem medida nula.*

Demonstração: Se f é descontínua num ponto $x \in X$, existe uma seqüência de pontos $x_k \in X$ tal que $(f(x_k))$ não converge para $f(x)$. Daí resulta que $\overline{f}$ também é descontínua no ponto x. Por sua vez, os pontos de descontinuidade de $\overline{f}$, ou já eram descontinuidades de f ou estão na fronteira de X. Assim, podemos escrever $D_f \subset D_{\overline{f}} \subset D_f \cup \partial X$. Como ∂X tem medida nula, vemos que med. $D_f = 0 \Leftrightarrow$ med. $D_{\overline{f}} = 0$. Ora, por definição, f integrável significa $\overline{f}$ integrável, ou seja, med. $D_f = 0$. Isto demonstra o teorema.

Mais geralmente, dada $f\colon X \to \mathbb{R}$ limitada no conjunto J-mensurável $X \subset \mathbb{R}^m$, definem-se a *integral inferior* e a *integral superior* de f no conjunto X pondo-se, respectivamente:

$$\underline{\int}_X f(x)dx = \underline{\int}_A \overline{f}(x)dx \quad \text{e} \quad \overline{\int}_X f(x)dx = \overline{\int}_A \overline{f}(x)dx,$$

onde A é um bloco de $\mathbb{R}^m$ contendo X e $\overline{f}\colon A \to \mathbb{R}$ é a extensão de f que se anula nos pontos de $A - X$. Obviamente, f é integrável em X se, e somente se, suas integrais (inferior e superior) sobre X coincidem.

Para a integral de uma função sobre um conjunto J-mensurável valem regras análogas às do Teorema 3, a saber:

Teorema 8. *Sejam $f, g\colon X \to \mathbb{R}$ funções integráveis no conjunto J-mensurável $X \subset \mathbb{R}^m$ e $c \in \mathbb{R}$. Então*

a) *As funções $c \cdot f$, $f + g\colon X \to \mathbb{R}$ são integráveis, com*

$$\int_X (c \cdot f)(x)dx = c\int_X f(x)dx \text{ e}$$

$$\int_X [f(x) + g(x)] = \int_X f(x)dx + \int_X g(x)dx.$$

b) *Se* $f(x) \leq g(x)$ *para todo* $x \in X$ *então*

$$\int_X f(x)dx \leq \int_X g(x)dx.$$

Em particular, se $m \leq f(x) \leq M$ *para todo* $x \in X$ *então*

$$m \cdot \text{vol. } X \leq \int_X f(x)dx \leq M \cdot \text{vol. } X.$$

c) *A função* $x \mapsto |f(x)|$ *é integrável e*

$$\left| \int_X f(x)dx \right| \leq \int_X |f(x)|dx.$$

Em particular, se $|f(x)| \leq K$ *para todo* $x \in X$ *então*

$$\left| \int_X f(x)dx \right| \leq K \cdot \text{vol. } X.$$

d) *Se* f *é contínua e* X *é conexo então existe* $x_0 \in X$ *tal que*

$$\int_X f(x)dx = f(x_0) \cdot \text{vol. } X.$$

Demonstração: a) Sejam $A \subset \mathbb{R}^m$ um bloco contendo X e $\overline{f}, \overline{g} \colon A \to \mathbb{R}$ as extensões de f e g, iguais a zero em $A - X$. As extensões de $c \cdot f$ e $f + g$ são respectivamente $c \cdot \overline{f}$ e $\overline{f} + \overline{g}$. Por hipótese, $\overline{f}$ e $\overline{g}$ são integráveis em A, logo o mesmo ocorre com $c \cdot \overline{f}$ e $\overline{f} + \overline{g}$. Portanto, $c \cdot f$ e $f + g$ são integráveis em X, com

$$\int_X (c \cdot f)(x)dx = \int_A c \cdot \overline{f}(x)dx = c \cdot \int_A \overline{f}(x)dx = c \cdot \int_X f(x)dx \text{ e}$$

$$\int_X [f(x) + g(x)]dx = \int_A [\overline{f}(x) + \overline{g}(x)]dx = \int_A \overline{f}(x)dx + \int_A \overline{g}(x)dx =$$

$$= \int_X f(x)dx + \int_X g(x)dx, \text{ em virtude do Teorema 3.}$$

b) Se $f(x) \leq g(x)$ para todo $x \in X$ então $\overline{f}(x) \leq \overline{g}(x)$ para todo $x \in A$. Logo

$$\int_X f(x)dx = \int_A \overline{f}(x)dx \leq \int_A \overline{g}(x)dx = \int_X g(x)dx.$$

Se tivermos $m \leq f(x) \leq M$ para todo $x \in M$ então $m \cdot \chi_X(x) \leq \overline{f}(x) \leq M \cdot \chi_X(x)$ para todo $x \in A$, logo

$$m \cdot \text{vol. } X = m \int_A \chi_X(x)dx = \int_A m \cdot \chi_X(x)dx \leq \int_A \overline{f}(x)dx = \\ = \int_X f(x)dx$$

e, analogamente,

$$\int_X f(x)dx \leq M \cdot \text{vol. } X.$$

c) Seja $g\colon X \to \mathbb{R}$ definida por $g(x) = |f(x)|$. Evidentemente $D_g \subset D_f$, logo g é integrável. A extensão $\overline{g}\colon A \to \mathbb{R}$ é dada por $\overline{g}(x) = |\overline{f}(x)|$, onde $\overline{f}\colon A \to \mathbb{R}$ é a extensão de f. Portanto

$$\left|\int_X f(x)dx\right| = \left|\int_A \overline{f}(x)dx\right| \leq \int_A |\overline{f}(x)|dx = \\ = \int_A \overline{g}(x)dx = \int_X g(x)dx = \int_X |f(x)|dx.$$

Se for $|f(x)| \leq K$ para todo $x \in X$ então

$$\left|\int_X f(x)dx\right| \leq \int_X |f(x)|dx \leq \int_X K \cdot dx = K \cdot \text{vol. } X.$$

d) Em virtude da conexidade de X, a imagem $f(X)$ é um intervalo, cujos extremos são m e M. Como $\dfrac{1}{\text{vol. } X}\displaystyle\int_X f(x)dx$ pertence a esse intervalo, é igual a $f(x_0)$ para algum $x_0 \in X$.

Teorema 9. *Sejam $X, Y \subset \mathbb{R}^m$ conjuntos J-mensuráveis. Uma função $f\colon X \cup Y \to \mathbb{R}$ é integrável se, e somente se, suas restrições $f|X$ e $f|Y$ são integráveis. No caso afirmativo,*

$$\int_{X\cup Y} f(x)dx + \int_{X\cap Y} f(x)dx = \int_X f(x)dx + \int_Y f(x)dx.$$

Em particular, se X e Y não possuem pontos interiores em comum, então

$$\int_{X\cup Y} f(x)dx = \int_X f(x)dx + \int_Y f(x)dx.$$

Demonstração: Se indicarmos com D, D_X e D_Y respectivamente os conjuntos dos pontos de descontinuidade de f, $f|X$ e $f|Y$, teremos $D_X \cup D_Y \subset D \subset D_X \cup D_Y \cup \partial X \cup \partial Y$. Sabemos que ∂X e ∂Y têm medida nula. Portanto med. $D = 0 \Leftrightarrow$ med. $D_X =$ med. $D_Y = 0$, ou seja, f é integrável se, e somente se, $f|X$ e $f|Y$ o são. No caso afirmativo, sejam A um bloco em $\mathbb{R}^m$ contendo $X \cup Y$ e $\overline{f}\colon A \to \mathbb{R}$ a extensão de f que se anula em $A - (X \cup Y)$. Então $\overline{f} = \overline{f} \cdot \chi_{X\cup Y}$ logo, da igualdade $\chi_{X\cup Y} + \chi_{X\cap Y} = \chi_X + \chi_Y$ concluimos $\overline{f} + \overline{f} \cdot \chi_{X\cap Y} = \overline{f} \cdot \chi_X + \overline{f} \cdot \chi_Y$

$$\int_{X\cup Y} f + \int_{X\cap Y} f = \int_A \overline{f} + \int_A \overline{f} \cdot \chi_{X\cap Y} =$$
$$= \int_A \overline{f} \cdot \chi_X + \int_A \overline{f} \cdot \chi_Y = \int_X f + \int_Y f.$$

Se X e Y não possuem pontos interiores em comum temos vol. $(X \cap Y) = 0$, em virtude do Exemplo 6. Existe $K > 0$ tal que $|f(x)| \le K$ para todo $x \in X$. Então $|\int_{X\cap Y} f(x)dx| \le K \cdot \text{vol.}\,(X \cap Y) = 0$, logo $\int_{X\cap Y} f(x)dx = 0$, daí resultando a última afirmação do teorema.

Corolário. *Seja $f\colon X \to \mathbb{R}$ integrável no conjunto J-mensurável $X \subset \mathbb{R}^m$. Se $Y \subset X$ é J-mensurável e $X - Y$ tem interior vazio então $\int_X f(x)dx = \int_Y f(x)dx$. Em particular, se $U = \text{int.}\,X$ então $\int_X f(x)dx = \int_U f(x)dx$.*

Isto resulta da última parte do teorema acima, aplicada à igualdade $X = (X - Y) \cup Y$, juntamente com as seguintes observações: primeiro, $X - Y$ é J-mensurável pois se tomarmos um bloco $A \supset X$ teremos $X - Y = X \cap (A - Y)$ e poderemos usar o Teorema 6. Segundo, como o conjunto J-mensurável $X - Y$ tem interior vazio, seu volume é zero (Exemplo 6) logo $\int_{X-Y} f(x)dx = 0$. Portanto

$$\int_X f(x)dx = \int_{X-Y} f(x)dx + \int_Y f(x)dx = \int_Y f(x)dx.$$

Finalmente, se $U = \text{int.}\,X$ então $\partial U \subset \partial X$, logo U é J-mensurável. Como $X - U = \partial X$ tem interior vazio, temos

$$\int_X f(x)dx = \int_U f(x)dx.$$

Observação: O teorema acima mostra que, ao considerarmos integrais sobre conjuntos J-mensuráveis em $\mathbb{R}^m$, não há perda de generalidade em

supor que tais conjuntos são abertos. Convém notar, entretanto, que nem todo subconjunto aberto limitado de $\mathbb{R}^m$ é J-mensurável. (Veja Exemplo 9 adiante.) No Capítulo VII, §6 estenderemos a noção de integral de modo que se possa definir a integral ("imprópria") de uma função sobre um aberto, mesmo que este não seja J-mensurável. Desde já, existe porém um caso particular em que podemos definir a integral de uma função $f\colon U \to \mathbb{R}$, definida num aberto $U \subset \mathbb{R}^m$, mesmo que U não seja J-mensurável. É quando f tem suporte compacto. [O *suporte* de f é $\text{supp}\cdot f$, fecho (relativamente a U) do conjunto dos pontos x tais que $f(x) \neq 0$.] Se $\text{supp}\cdot f$ é compacto, podemos definir $\int_U f(x)dx = \int_K f(x)dx$, onde K é qualquer conjunto J-mensurável, tal que $\text{supp}\cdot f \subset K \subset U$. A integral existe desde que o conjunto dos pontos de descontinuidade de f tenha medida nula. O compacto J-mensurável K também existe: na norma do máximo, observar que $d(\text{supp}\cdot f, \mathbb{R}^m - U) = \delta > 0$, cobrir o compacto $\text{supp}\cdot f$ com um número finito de cubos de arestas $< \delta$ e tomar K igual à reunião desses cubos. A integral $\int_K f(x)dx$, evidentemente, não vai depender do conjunto K escolhido nessas condições pois f se anula fora de $\text{supp}\cdot f$.

4 A integral como limite de somas de Riemann

Estenderemos agora o conceito de partição, a fim de exprimir a integral como limite de somas de Riemann, nas quais o domínio da função é decomposto como reunião de conjuntos J-mensuráveis mais gerais do que blocos. Essa formulação da idéia de integral é a mais comum nas exposições intuitivas do Cálculo, principalmente as que visam as aplicações.

Seja $X \subset \mathbb{R}^m$ um conjunto J-mensurável. Uma *decomposição* de X é uma coleção finita $D = \{X_1, \ldots, X_k\}$ de conjuntos J-mensuráveis tais que $X = X_1 \cup \cdots \cup X_k$ e $X_i \cap X_j$ tem interior vazio se $i \neq j$. A *norma* da decomposição D é o número $|D|$ = maior diâmetro dos conjuntos $X_i \in D$. Uma *decomposição pontilhada* é um par $D^* = (D, (\xi_i))$ onde $\xi_1 \in X_1, \ldots, \xi_k \in X_k$. Dadas a decomposição $D = \{X_1, \ldots, X_k\}$ de X e a função limitada $f\colon X \to \mathbb{R}$, escreveremos, para cada $i = 1, \ldots, k$, $m_i = m_i(f) = \inf.\{f(x); x \in X_i\}$, $M_i = M_i(f) = \sup.\{f(x); x \in X_i\}$ e definiremos a *soma inferior* e a *soma superior* de f relativamente à

decomposição D pondo, respectivamente:

$$s(f;D) = \sum_{i=1}^{k} m_i \cdot \text{vol. } X_i \quad \text{e} \quad S(f;D) = \sum_{i=1}^{k} M_i \cdot \text{vol. } X_i \,.$$

A *soma de Riemann* de f relativamente à decomposição pontilhada $D^* = (D, (\xi_i))$ é definida por

$$\Sigma(f;D^*) = \sum_{i=1}^{k} f(\xi_i) \cdot \text{vol. } X_i \,.$$

Um exemplo de decomposição é fornecido por uma partição P de um bloco A. Neste caso, os conjuntos X_i da decomposição são os blocos $B \in P$.

Seja $f\colon X \to \mathbb{R}$ integrável no conjunto J-mensurável $X \subset \mathbb{R}^m$. Nosso objetivo é provar que $\int_X f(x)dx$ é o limite das somas de Riemann $\Sigma(f;D^*)$ quando a norma $|D|$ tende para zero. Isto significa que, para todo $\varepsilon > 0$ dado, pode-se obter $\delta > 0$ tal que $|\int_X f(x)dx - \Sigma(f;D^*)| < \varepsilon$, seja qual for a decomposição D de X com $|D| < \delta$ e seja qual for a maneira D^* de pontilhar essa decomposição. Mostraremos também que vale a recíproca: se existe $\lim_{|D|\to 0} \Sigma(f;D^*) = c$ então f é integrável sobre X e $\int_X f(x)dx = c$.

Nossa demonstração fará uso do lema seguinte. Lembremos que, se $X, Y \subset \mathbb{R}^m$ são subconjuntos não-vazios e $\delta > 0$ então a afirmação $d(X,Y) < \delta$ significa que existem $x \in X$ e $y \in Y$ tais que $|x - y| < \delta$.

Lema. *Sejam $Y \subset X \subset \mathbb{R}^m$ J-mensuráveis, com* vol. $Y = 0$. *Para todo $\varepsilon > 0$ dado, existe $\delta > 0$ tal que se D é qualquer decomposição de X com $|D| < \delta$ então a soma dos volumes dos conjuntos $X_i \in D$ tais que $d(X_i, Y) < \delta$ é menor do que ε.*

Demonstração: Dado $\varepsilon > 0$ podemos cobrir Y com uma coleção finita de blocos B cuja soma dos volumes é menor do que ε. Tomando arbitrariamente $\delta > 0$, ponhamos cada um desses blocos $B = \Pi[a_i, b_i]$ dentro do bloco $B' = \Pi[a_i - 2\delta, b_i + 2\delta]$. Evidentemente, $\lim_{\delta\to 0} \text{vol. } B' = \text{vol. } B$. Logo existe $\delta > 0$ tal que a soma dos volumes dos blocos B' ainda é menor do que ε. Ora, se Z é um conjunto de diâmetro $< \delta$ tal que $d(Z, B) < \delta$ então $Z \subset B'$. Portanto, se $D = \{X_1, \ldots, X_k\}$ é uma decomposição de X com $|D| < \delta$, vemos que $d(X_i; Y) < \delta \Rightarrow d(X_i; B) < \delta$ para algum

$B \Rightarrow X_i \subset B'$. Assim, a soma dos volumes dos conjuntos $X_i \in D$ tais que $d(X_i, Y) < \delta$ não excede a soma dos volumes dos blocos B', logo é menor do que ε.

Teorema 10. *Para toda função $f\colon X \to \mathbb{R}$, limitada no conjunto J-mensurável $X \subset \mathbb{R}^m$, tem-se*

$$\underline{\int}_X f(x)dx = \lim_{|D|\to 0} s(f;D) \quad e \quad \overline{\int}_X f(x)dx = \lim_{|D|\to 0} S(f;D).$$

Demonstração: Basta provar a segunda afirmação. Se somarmos uma constante c à função f, tanto a integral superior como o limite acima serão aumentados de $c \cdot \text{vol. } X$. Logo podemos supor, sem perda de generalidade, que $0 \leq f(x) \leq K$ para todo $x \in X$. Sejam A um bloco em $\mathbb{R}^m$, contendo X, e $\overline{f}\colon A \to \mathbb{R}$ a extensão de f tal que $\overline{f}(x) = 0$ se $x \in A - X$. Queremos provar que, para todo $\varepsilon > 0$ dado, podemos obter $\delta > 0$ tal que, se D é qualquer decomposição de X com $|D| < \delta$ então $|S(f;D) - \overline{\int}_X f(x)dx| < \varepsilon$. Ora, dado $\varepsilon > 0$, existe uma partição P_0 de A tal que

$$S(\overline{f};P_0) < \overline{\int}_X f(x)dx + \varepsilon/2.$$

Seja Y a reunião das faces próprias dos blocos de P_0. Temos vol. $Y = 0$. Pelo lema, existe $\delta > 0$ tal que, para toda decomposição D de X com $|D| < \delta$, a soma dos volumes dos conjuntos $X_i \in D$ tais que $d(X_i, Y) < \delta$ é menor do que $\varepsilon/2K$. Seja, portanto, D uma decomposição de X cuja norma $|D|$ é menor do que δ. Chamemos de X_α os conjuntos da decomposição D tais que $d(X_\alpha, Y) < \delta$. Os demais conjuntos de D serão chamados de X_β. Note que cada X_β deve estar contido no interior de algum bloco da partição P_0. [Do contrário existiriam pontos $x, y \in X_\beta$ em blocos distintos da partição, logo o segmento de reta $[x, y]$ conteria algum ponto de Y. Como $|x - y| < \delta$, isto daria $d(X_\beta, Y) < \delta$, um absurdo.] Pondo $M_\alpha = \sup.\{f(x); x \in X_\alpha\}$ e $M_\beta = \sup.\{f(x); x \in X_\beta\}$, temos:

$$S(f;D) = \Sigma\, M_\alpha \cdot \text{vol. } X_\alpha + \Sigma\, M_\beta \cdot \text{vol. } X_\beta\,,$$

onde $\Sigma\, M_\alpha \cdot \text{vol. } X_\alpha \le K \cdot \Sigma\, \text{vol. } X_\alpha < \varepsilon/2$ e

$$\Sigma\, M_\beta \cdot \text{vol. } X_\beta = \sum_{B \in P_0} \left(\sum_{X_\beta \subset B} M_\beta \cdot \text{vol. } X_\beta \right) \le \sum_{B \in P_0} M_\beta \cdot \text{vol. } B =$$
$$= S(\overline{f}; P_0) < \overline{\int}_X f(x)dx + \varepsilon/2.$$

Assim, $|D| < \delta \Rightarrow S(f;D) < \overline{\int}_X f(x)dx + \varepsilon$. Mostraremos a seguir que, para toda decomposição D de X, temos $S(f;D) \ge \overline{\int}_X f(x)dx$. Com efeito, seja Z a reunião das fronteiras dos conjuntos X_i da decomposição D. Temos vol. $Z = 0$. Pelo lema, existe $\delta' > 0$ tal que, para toda partição P do bloco A com $|P| < \delta'$, a soma dos volumes dos blocos de P que intersectam Z é menor do que ε/K. Tomando $|P| < \delta'$, temos $S(\overline{f}; P) = \Sigma\, M_\beta \cdot \text{vol. } B + \Sigma\, M_C \cdot \text{vol. } C$, onde chamamos de B os blocos de P que intersectam Z e de C os que estão contidos no interior de algum $X_i \in D$. [Observe que, sendo conexo, um bloco que não esteja contido no interior de algum X_i deve intersectar Z, pelo Teorema da Alfândega.] Ora,

$$\Sigma\, M_\beta \cdot \text{vol. } B \le K \cdot \Sigma\, \text{vol. } B < \varepsilon \quad \text{e}$$

$$\Sigma\, M_C \cdot \text{vol. } C = \sum_i \left(\sum_{C \subset X_i} M_C \cdot \text{vol. } C \right) \le \sum_i M_i \left(\sum_{C \subset X_i} \text{vol. } C \right) \le$$
$$\le \sum_i M_i \cdot \text{vol. } X_i = S(f;D).$$

Logo $\overline{\int}_X f(x)dx \le S(\overline{f}; P) < S(f;D) + \varepsilon$.
Como $\varepsilon > 0$ é arbitrário, segue-se que $\overline{\int}_X f(x)dx \le S(f;D)$ para toda decomposição D de X. Isto conclui a demonstração.

Corolário (da demonstração). *Se $f\colon X \to \mathbb{R}$ é limitada e não-negativa no conjunto J-mensurável $X \subset \mathbb{R}^m$ então*

$$\underline{\int}_X f(x)dx = \sup_D .s(f;D) \quad e \quad \overline{\int}_X f(x)dx = \inf_D .S(f;D).$$

Teorema 11. *Seja $f\colon X \to \mathbb{R}$ uma função limitada no conjunto J-mensurável $X \subset \mathbb{R}^m$. A fim de que f seja integrável, é necessário e suficiente que exista o limite $I = \lim\limits_{|D|\to 0} \Sigma(f;D^*)$. No caso afirmativo, este limite é igual à integral de f sobre X.*

Demonstração: Para toda decomposição pontilhada D^* de X tem-se, evidentemente,

$$s(f;D) \leq \Sigma(f;D^*) \leq S(f;D).$$

Pelo teorema anterior, se f é integrável, vale

$$\lim_{|D|\to 0} s(f;D) = \lim_{|D|\to 0} S(f;D) = \int_X f(x)dx.$$

Logo, quando f é integrável, existe o limite

$$\lim_{|D|\to 0} \Sigma(f;D^*) = \int_X f(x)dx.$$

Reciprocamente, suponhamos que exista o limite I do enunciado. Para provar que f é integrável podemos, sem perda de generalidade, supor que $f \geq 0$. Dado arbitrariamente $\varepsilon > 0$, existe uma decomposição $D = \{X_1, \ldots, X_k\}$ de X tal que $|\Sigma(f;D^*) - I| < \varepsilon/4$ seja qual for a maneira de pontilhar D. Fixemos D e a pontilhemos de duas maneiras. Em primeiro lugar, podemos escolher em cada conjunto $X_i \in D$ um ponto ξ_i tal que $f(\xi_i) < m_i + \varepsilon/(4k \cdot \text{vol. } X_i)$. Isto nos dá uma decomposição pontilhada D^* com

$$\Sigma(f;D^*) = \sum_{i=1}^{k} f(\xi_i)\cdot \text{vol. } X_i < \sum_{i=1}^{k} m_i \cdot \text{vol. } X_i + \varepsilon/4 =$$
$$= s(f;D) + \varepsilon/4.$$

De modo análogo, obtemos uma decomposição pontilhada $D^\#$ tal que $S(f;D) - \varepsilon/4 < \Sigma(f;D^\#)$. Desta maneira,

$$\Sigma(f;D^*) - \frac{\varepsilon}{4} < s(f;D) \leq S(f;D) < \Sigma(f;D^\#) + \frac{\varepsilon}{4}\cdot$$

Como os números $\Sigma(f;D^*)$ e $\Sigma(f;D^\#)$ pertencem ao intervalo $I - \varepsilon/4, I + \varepsilon/4)$, segue-se que $s(f;D)$ e $S(f;D)$ pertencem ao intervalo

$(I - \varepsilon/2, I + \varepsilon/2)$ e portanto $S(f;D) - s(f;D) < \varepsilon$. Sendo $\varepsilon > 0$ arbitrário, temos

$$\underline{\int}_X f(x)dx = \sup_D .s(f;D) = \inf_D .S(f;D) = \overline{\int}_X f(x)dx,$$

em virtude do corolário do teorema anterior. Logo f é integrável e sua integral é igual a I.

Lema de Duhamel. *Seja $f\colon X \to \mathbb{R}$ integrável no conjunto J-mensurável $X \subset \mathbb{R}^m$. Para cada decomposição $D = \{X_1, \ldots, X_k\}$ de X, suponhamos dados números $\eta_1 = \eta_1(D), \ldots, \eta_k = \eta_k(D)$ tais que $\lim_{|D|\to 0} \eta_i = 0$. [Isto significa que, para qualquer $\varepsilon > 0$ dado, pode-se obter $\delta > 0$ tal que $|D| < \delta \Rightarrow |\eta_1(D) < \varepsilon, \ldots, |\eta_k(D)| < \varepsilon$.] Nestas condições, tem-se*

$$\lim_{|D|\to 0} \Sigma[f(\xi_i) + \eta_i] \operatorname{vol.} X_i = \int_X f(x)dx.$$

Demonstração: Basta observar que, se $\eta_i < \varepsilon / \operatorname{vol.} X$ para $i = 1, \ldots, k$, então $\Sigma\, \eta_i \cdot \operatorname{vol.} X_i < \dfrac{\varepsilon}{\operatorname{vol.} X} \cdot \Sigma \operatorname{vol.} X_i = \varepsilon$, logo $\lim_{|D|\to 0} \Sigma\, \eta_i \cdot \operatorname{vol.} X_i = 0$. Portanto $\lim \Sigma[f(\xi_i) + \eta_i] \cdot \operatorname{vol.} X_i = \lim \Sigma\, f(\xi_i) \operatorname{vol.} X_i + \lim \Sigma\, \eta_i \cdot \operatorname{vol.} X_i = \lim \Sigma\, f(\xi_i) \cdot \operatorname{vol.} X_i = \int_X f(x)dx$.

5 Integração repetida

A redução de uma integral sobre um bloco m-dimensional ("integral múltipla") a uma seqüência de m integrais de funções de uma variável ("integral repetida") é um eficaz instrumento de cálculo, além de ter importância teórica, como constataremos, por exemplo, na demonstração do Teorema de Stokes, no Capítulo VII. Tal redução será provada agora. Dela decorre, em particular, o Teorema da Inversão de Ordem nas Integrais Repetidas, que demonstramos no Capítulo III, §6.

Evidentemente, para reduzir uma integral sobre um bloco a sucessivas integrais sobre intervalos basta considerar $A = A_1 \times A_2 \subset \mathbb{R}^{m+n}$, onde $A_1 \subset \mathbb{R}^m$ e $A_2 \subset \mathbb{R}^n$ são blocos, e mostrar que toda integral sobre A se obtém integrando-se primeiro sobre A_2 e depois sobre A_1 (ou vice-versa).

É o que faremos. Os pontos de $A_1 \times A_2$ serão escritos como (x, y), onde $x \in A_1$, $y \in A_2$ e se $f\colon A_1 \times A_2 \to \mathbb{R}$ é integrável, sua integral será indicada com a notação $\int_{A_1 \times A_2} f(x, y)dxdy$.

Dada $f\colon A_1 \times A_2 \to \mathbb{R}$, para cada $x \in A_1$ escreveremos $f_x\colon A_2 \to \mathbb{R}$, onde $f_x(y) = f(x, y)$, $y \in A_2$. Portanto f_x é essencialmente a restrição de f ao bloco n-dimensional $\{x\} \times A_2$. Mesmo supondo f integrável, a função f_x pode, para alguns valores de $x \in A_1$, não ser integrável. Com efeito, o conjunto dos pontos de descontinuidade de f tem medida nula em $\mathbb{R}^{m+n}$ mas sua interseção com algum bloco $\{x\} \times A_2$ pode não ter medida n-dimensional nula.

Exemplo 7. Seja $f\colon [0,1] \times [0,1] \to \mathbb{R}$ definida por $f(x,y) = 0$ se $x \neq 1/2$, enquanto $f(1/2, y) = 1$ se y é racional e $f(1/2, y) = 0$ caso y seja irracional. O conjunto dos pontos de descontinuidade de f é o segmento vertical $\{1/2\} \times [0,1]$, que tem medida nula em $\mathbb{R}^2$. Assim, f é integrável. Para todo $x \neq 1/2$, $f_x\colon [0,1] \to \mathbb{R}$ é identicamente nula e portanto integrável. Mas $f_{1/2}$ é descontínua em todos os pontos do intervalo $[0,1]$, logo não é integrável.

Observação: Pode-se mostrar (v., por exemplo, o Exerc. 5.3, adiante) que se f é integrável então o conjunto dos pontos $x \in A_1$ tais que f_x não é integrável tem medida nula. Isto é parte do importante "Teorema de Fubini", o qual contém também o teorema que provaremos abaixo.

Teorema da Integração Repetida. *Seja $f\colon A_1 \times A_2 \to \mathbb{R}$ integrável no produto dos blocos $A_1 \subset \mathbb{R}^m$ e $A_2 \subset \mathbb{R}^n$. Para todo $x \in A_1$, seja $f_x\colon A_2 \to \mathbb{R}$ definida por $f_x(y) = f(x,y)$ e ponhamos*

$$\varphi(x) = \underline{\int}_{A_2} f_x(y)dy, \quad \psi(x) = \overline{\int}_{A_2} f_x(y)dy.$$

As funções $\varphi, \psi\colon A_1 \to \mathbb{R}$, assim definidas, são integráveis, com

$$\int_{A_1} \varphi(x)dx = \int_{A_1} \psi(x)dx = \int_{A_1 \times A_2} f(x,y)dxdy,$$

isto é:

$$\int_{A_1 \times A_2} f(x,y)dxdy = \int_{A_1} dx \left(\underline{\int}_{A_2} f(x,y)dy \right) = \int_{A_1} dx \left(\overline{\int}_{A_2} f(x,y)dy \right).$$

Demonstração: Seja $P = P_1 \times P_2$ uma partição qualquer de $A_1 \times A_2$. Os blocos de P são os produtos $B_1 \times B_2$, onde $B_1 \in P_1$ e $B_2 \in P_2$. A soma inferior de f relativa à partição P se escreve

$$s(f;P) = \Sigma\, m_{B_1 \times B_2} \cdot \text{vol. } B_1 \cdot \text{vol. } B_2 =$$
$$= \sum_{B_1 \in P_1} \left(\sum_{B_2 \in P_2} m_{B_1 \times B_2} \cdot \text{vol. } B_2 \right) \text{vol. } B_1\,.$$

Para todo $x \in B_1$, $\;m_{B_1 \times B_2} = m_{B_1 \times B_2}(f) \leq m_{B_2}(f_x)$, logo

$$\sum_{B_2 \in P_2} m_{B_1 \times B_2} \cdot \text{vol. } B_2 \leq \sum_{B_2 \in P_2} m_{B_2}(f_x) \cdot \text{vol. } B_2 \leq \varphi(x).$$

Como esta desigualdade vale para todo $x \in B_1$, concluimos que

$$\sum_{B_2 \in P_2} m_{B_1 \times B_2} \cdot \text{vol. } B_2 \leq m_{B_1}(\varphi). \text{ Daí}$$

$$s(f;P) \leq \sum_{B_1 \in P_1} m_{B_1}(\varphi) \cdot \text{vol. } B_1 = s(\varphi; P_1).$$

Analogamente se prova a desigualdade $S(\varphi;P_1) \leq S(f;P)$. Logo

$$s(f;P) \leq s(\varphi;P_1) \leq S(\varphi;P_1) \leq S(f;P),$$

para qualquer partição $P = P_1 \times P_2$. Como f é integrável, decorre imediatamente que φ é integrável e que $\int_{A_1} \varphi(x)dx = \int_{A_1 \times A_2} f(x,y)dxdy$. A afirmação sobre ψ se prova da mesma maneira.

Corolário. *Se $f\colon A_1 \times A_2 \to \mathbb{R}$ é integrável então*

$$\int_{A_1} dx \left(\overline{\int}_{A_2} f(x,y)dy \right) = \int_{A_2} dy \left(\overline{\int}_{A_1} f(x,y)dx \right) =$$
$$= \int_{A_1 \times A_2} f(x,y)dxdy$$

e valem mais 3 igualdades análogas, que se obtêm tomando integrais superiores e inferiores dentro dos parênteses. Em particular, se f_x e f_y são contínuas para quaisquer $x \in A_1$ e $y \in A_2$ (por exemplo, se f é

contínua) então

$$\int_{A_1\times A_2} f(x,y)dxdy = \int_{A_1} dx \left(\int_{A_2} f(x,y)dy\right) = \\ = \int_{A_2} dy \left(\int_{A_1} f(x,y)dx\right).$$

Com efeito, tudo que se fez com x no teorema acima pode ser feito também com y.

Exemplo 8. Seja $X \subset [0,1]$ um conjunto de Cantor de medida positiva, como no Exemplo 3. Afirmamos que o conjunto $Y = X \times [0,1]$ não tem medida nula em $\mathbb{R}^2$. Com efeito, o quadrado $A = [0,1] \times [0,1]$ contém Y e a função característica $\chi_Y \colon A \to \mathbb{R}$ é contínua em todos os pontos de $A - Y$. Se Y tivesse medida nula então $f = \chi_Y$ seria integrável. Com a notação do teorema anterior, para todo $x \in [0,1]$ temos f_x constante (igual a zero se $x \notin X$, igual a 1 para x em X). Como $\varphi(x) = \int_0^1 f_x(y)dy$, segue-se que $\varphi \colon [0,1] \to \mathbb{R}$ é a função característica de X. Pelo referido teorema, φ seria integrável, o que equivale a dizer que o conjunto X seria J-mensurável, o que é absurdo pois med. $\partial X =$ med. $X \neq 0$.

Exemplo 9. Usando o conjunto Y do exemplo acima, exibiremos um aberto limitado $U \subset \mathbb{R}^2$ que não é J-mensurável. Basta tomar qualquer retângulo aberto V em $\mathbb{R}^2$, contendo Y, e pôr $U = V - Y$. Então $\partial U \supset Y$, logo ∂U não tem medida nula, isto é, U não é J-mensurável.

6 Mudança de variáveis

A fórmula de mudança de variáveis para integrais simples é quase automática: na integral $\int_{h(a)}^{h(b)} f(y)dy$, se $h \colon [a,b] \to \mathbb{R}$ é de classe C^1, põe-se $y = h(x)$ e $dy = h'(x)dx$. Quando x varia entre a e b, $h(x)$ vai de $h(a)$ até $h(b)$ e tem-se:

$$\int_{h(a)}^{h(b)} f(y)dy = \int_a^b f(h(x))h'(x)dx.$$

Isto foi demonstrado, para f contínua, (sem supor h injetiva) no Vol. 1, pág. 326, como uma simples conseqüência do Teorema Fundamental do Cálculo.

Para integrais múltiplas, a fórmula de mudança de variáveis se torna

$$\int_{h(X)} f(y)dy = \int_X f(h(x)) \cdot |\,\text{det. } h'(x)|dx.$$

Aqui, supõe-se que h seja um difeomorfismo de classe C^1 numa vizinhança de X, mas $f\colon h(X) \to \mathbb{R}$ é apenas integrável. Esta fórmula é o resultado mais importante deste capítulo.

Comparando as regras de mudança de variáveis para integrais simples e múltiplas, acharemos natural que o determinante jacobiano substitua a derivada, mas dificilmente deixaremos de estranhar o valor absoluto. Por que ele não ocorre no caso de uma variável? A resposta é: ele ocorre, sim, mas fica oculto por uma notação conveniente. (E não poderia deixar de ocorrer, pois a definição que demos no §1 deste capítulo para a integral de uma função de m variáveis inclui o caso particular $m = 1$.)

Tudo começa com a convenção $\int_a^b f(x)dx = -\int_b^a f(x)dx$, que faz da integral simples uma integral "orientada". Dado o intervalo $I = [a, b]$, com $a < b$, indiquemos com a notação $\int_I f(x)dx$ a integral da função $f\colon I \to \mathbb{R}$, definida como no §1, já que I é um bloco de dimensão 1. Então $\int_I f(x)dx = \int_a^b f(x)dx = -\int_b^a f(x)dx$. Se h é um difeomorfismo de classe C^1 de I sobre o intervalo $J = h(I)$, ou se tem $h'(x) > 0$ para todo $x \in [a, b]$, e então $h(a) < h(b)$, ou vale $h'(x) < 0$ para todo x, donde $h(b) < h(a)$. No segundo caso, pela fórmula provada no Vol. 1, tem-se

$$\int_{h(I)} f(y)dy = -\int_{h(a)}^{(h(b)} f(y)dy = \int_a^b f(h(x))(-h'(x))dx =$$
$$= \int_I f(h(x)) \cdot |h'(x)|dx,$$

fórmula que, evidentemente, também é válida no primeiro caso.

Observação: O estudo de integrais múltiplas orientadas será feito no próximo capítulo, onde veremos que os integrandos mais naturais nesse caso são as formas diferenciais.

O plano para a demonstração do Teorema de Mudança de Variáveis é o seguinte: primeiro provaremos uma fórmula para mudança linear de variável numa integral superior sobre um intervalo, depois demonstraremos o teorema no caso particular em que o difeomorfismo h é uma

transformação linear invertível. Usaremos o caso linear para provar, então, o teorema geral.

Lema 1. *Sejam* $T\colon \mathbb{R} \to \mathbb{R}$ *definida por* $T(x) = \alpha x + \beta$*, com* $\alpha \neq 0$*, e* $I \subset \mathbb{R}$ *um intervalo compacto. Dada* $f\colon J \to \mathbb{R}$*, limitada no intervalo* $J = T(I)$*, tem-se*

$$\overline{\int_J} f(y)dy = |\alpha| \overline{\int_I} f(\alpha x + \beta)dx.$$

Demonstração: Observe que as partições $Q = \{t_0, \ldots, t_k\}$ de J são todas do tipo $Q = \alpha \cdot P + \beta = \{\alpha \cdot s_0 + \beta, \ldots, \alpha \cdot s_k + \beta\}$, onde $P = \{s_0, \ldots, s_k\}$ é uma partição de I se $\alpha > 0$. Se for $\alpha < 0$ então a partição P tem a forma $P = \{s_k, \ldots, s_1, s_0\}$. Logo

$$\begin{aligned} S(f;Q) &= \Sigma\, M_i(f)(t_i - t_{i-1}) = \alpha \cdot \Sigma\, M_i(f \circ T)(s_i - s_{i-1}) = \\ &= -\alpha \cdot \Sigma\, M_i(f \circ T)(s_{i-1} - s_i), \text{ ou seja} \end{aligned}$$

$$S(f;Q) = \left\{ \begin{array}{c} \alpha \cdot S(f \circ T; P) \text{ se } \alpha > 0 \\ -\alpha \cdot S(f \circ T; P) \text{ se } \alpha < 0 \end{array} \right\} = |\alpha| \cdot S(f \circ T; P).$$

Segue-se imediatamente que

$$\overline{\int_J} f(y)dy = |\alpha| \overline{\int_I} f(\alpha x + \beta)dx.$$

Corolário. *Sejam* $f\colon [a,b] \to \mathbb{R}$ *uma função limitada e* $T\colon \mathbb{R} \to \mathbb{R}$ *definida por* $T(x) = \alpha x + \beta$*, com* $\alpha \neq 0$*. Suponha que* f *se anula fora de um conjunto arbitrário* $Y \subset [a,b]$ *e fora de* $T(Y) \subset [a,b]$*. Então*

$$\overline{\int_a^b} f(y)dy = \overline{\int_a^b} f(\alpha x + \beta)|\alpha|dx.$$

Com efeito, tomemos um intervalo compacto I, tal que $Y \subset I \subset [a,b]$ e $J = T(I) \subset [a,b]$. Usando o Lema 1, temos

$$\begin{aligned} \overline{\int_a^b} f(y)dy = \overline{\int_J} f(y)dy = \overline{\int_I} f(\alpha x + \beta)|\alpha|dx = \\ = \overline{\int_a^b} f(\alpha x + \beta)|\alpha|dx. \end{aligned}$$

Antes de passarmos ao próximo resultado observemos, de uma vez por todas, que se $h\colon U \to V$ é um difeomorfismo de classe C^1 entre abertos de $\mathbb{R}^m$ então (por h ser um homeomorfismo) tem-se $h(\partial X) = \partial h(X)$ para todo conjunto X tal que $\overline{X} \subset U$. Além disso, como h é localmente Lipschitziana, se $Y \subset U$ tem medida nula então $h(Y)$ tem medida nula. Assim, se X é J-mensurável e $\overline{X} \subset U$, sua imagem $h(X)$ é J-mensurável. Finalmente, se $f\colon h(X) \to \mathbb{R}$ então os conjuntos dos pontos de descontinuidade de f e $f \circ h\colon X \to \mathbb{R}$ estão relacionados pela igualdade $D_{f\circ h} = h^{-1}(D_f)$. Como h^{-1} também é localmente Lipschitziana, segue-se que med. $D_{f\circ h} = 0 \Leftrightarrow$ med. $D_f = 0$, isto é, f é integrável se, e somente se, $f \circ h$ é integrável.

Caso Linear da Mudança de Variáveis. *Sejam $T\colon \mathbb{R}^m \to \mathbb{R}^m$ uma transformação linear invertível, $X \subset \mathbb{R}^m$ um conjunto J-mensurável e $f\colon T(X) \to \mathbb{R}$ uma função integrável. Então*

$$\int_{T(X)} f(y)dy = \int_X f(T \cdot x)|\operatorname{det.} T|dx.$$

Demonstração: A observação básica é a seguinte: se a igualdade acima vale com $T = T_1$ e com $T = T_2$, então vale com $T = T_1T_2$. Ora, toda transformação linear invertível $T\colon \mathbb{R}^m \to \mathbb{R}^m$ se exprime como produto de "transformações elementares"(necessariamente invertíveis) dos dois tipos abaixo. (Veja a prova no Apêndice a este parágrafo.)

Tipo 1. $x = (x_1, x_2, \ldots, x_m) \mapsto T \cdot x = (\varphi(x), x_2, \ldots, x_m)$, com

$$\varphi(x) = \sum_{i=1}^{m} \alpha_i x_i\,.$$

Tipo 2. $x = (x_1, \ldots, x_i, \ldots, x_j, \ldots, x_m) \mapsto T \cdot x =$
$= (x_1, \ldots, x_j, \ldots, x_i, \ldots, x_m).$

Basta, portanto, considerar os casos em que T é de um desses tipos. Seja, primeiro, T do Tipo 1. Escrevamos $\varphi(x) = \alpha_1 \cdot x_1 + \beta$. Como T é invertível, tem-se $\alpha_1 \neq 0$. Seja $A = [a, b] \times A'$ um bloco m-dimensional contendo X e $T(X)$, onde A' é um bloco em R^{m-1}. Para simplificar a notação, consideraremos f definida em A, igual a zero em $A - T(X)$. Usando a notação $y = (t, w) \in A$ para os pontos de $A = [a, b] \times A'$, vemos que todo $y = (t, w) \in T(X)$ é da forma $y = T(x) = (\varphi(x), w) = (\alpha_1 s + \beta, w)$, onde $x = (s, w)$. Usando o Teorema da Integração Repetida

e o Corolário do Lema 1, (no qual, para cada $w \in A'$, tomamos $Y = Y_w = \{t \in [a,b]; (t,w) \in X\}$), temos

$$\begin{aligned}\int_{T(X)} f(y)dy &= \int_A f(y)dy = \int_{A'} dw \overline{\int_a^b} f_w(t)dt = \\ &= \int_{A'} dw \overline{\int_a^b} f_w(\alpha_1 s + \beta)|\alpha_1| ds = \\ &= \int_{A'} dw \overline{\int_a^b} f(\alpha_1 s + \beta, w) \cdot |\operatorname{det.} T| ds = \\ &= \int_A f(\alpha_1 s + \beta, w) \cdot |\operatorname{det.} T| ds dw = \\ &= \int_A f(T \cdot x) \cdot |\operatorname{det.} T| dx.\end{aligned}$$

Suponhamos agora que $T\colon \mathbb{R}^m \to \mathbb{R}^m$ seja uma transformação elementar do Tipo 2. Então det. $T = -1$, logo $|\operatorname{det.} T| = 1$. Devemos, portanto mostrar que $\int_{T(X)} f(y)dy = \int_X f(T \cdot x)dx$ quando T é do Tipo 2. Ora, para todo bloco

$$B = [a_1, b_1] \times \cdots \times [a_i, b_i] \times \cdots \times [a_j, b_j] \times \cdots \times [a_m, b_m],$$

temos $T(B) = [a_1, b_1] \times \cdots \times [a_j, b_j] \times \cdots \times [a_i, b_i] \times \cdots \times [a_m, b_m]$, logo $T(B)$ é um bloco cujo volume é igual ao de B. Como o volume de um conjunto J-mensurável $Z \subset \mathbb{R}^m$ é o ínfimo dos números $\Sigma \operatorname{vol.} B_i$, onde os B_i são blocos, com $Z \subset B_1 \cup \cdots \cup B_k$, segue-se que vol. $T(Z) =$ vol. Z para todo conjunto J-mensurável $Z \subset \mathbb{R}^m$. A integral $\int_{T(X)} f(y)dy$ pode ser calculada por meio do Teorema 11. Toda decomposição $T(X) = Y_1 \cup \cdots \cup Y_k$ é tal que $Y_i = T(X_i)$, onde $X = X_1 \cup \cdots \cup X_k$ é uma decomposição de X. Se $\xi_i \in Y_i$ então $\xi_i = T(\eta_i)$, $\eta_i \in X_i$. Logo

$$\begin{aligned}\int_{T(X)} f(y)dy &= \lim_{|D|\to 0} \Sigma f(\xi_i) \operatorname{vol.} Y_i = \lim_{|D|\to 0} \Sigma f(T \cdot \eta_i) \operatorname{vol.} T(X_i) = \\ &= \lim_{|D|\to 0} \Sigma f(T \cdot \eta_i) \cdot \operatorname{vol.} X_i = \int_X f(T \cdot x) \cdot dx,\end{aligned}$$

o que prova o teorema.

Corolário. *Seja $X \subset \mathbb{R}^m$ um conjunto J-mensurável. Para toda transformação linear $T\colon \mathbb{R}^m \to \mathbb{R}^m$, tem-se*

$$\operatorname{vol.} T(X) = |\operatorname{det.} T| \cdot \operatorname{vol.} X.$$

Com efeito, se T não é invertível então $T(X)$ está contida na imagem de T, que é uma superfície de dimensão $< m$ em $\mathbb{R}^m$, logo $T(X)$ tem interior vazio, e daí vol. $T(X) = 0 = |\det.\ T| \cdot \text{vol. } X$, pois, neste caso, det. $T = 0$. Se, porém, T é invertível, aplicamos o teorema acima, com $f = 1\colon T(X) \to \mathbb{R}$ a função constante, igual a 1, e temos

$$\text{vol. } T(X) = \int_{T(X)} 1 \cdot dx = \int_X 1 \cdot |\det.\ T| dx = \\ = |\det.\ T| \int_X 1 \cdot dx = |\det.\ T| \text{ vol. } X.$$

Exemplo 10. Dados $v_1, \ldots, v_m \in \mathbb{R}^m$, indicaremos (agora e no que se segue) com a notação $[v_1, \ldots, v_m]$ o *paralelepípedo* gerado pelos vetores dados, o qual é, por definição, o conjunto dos pontos de $\mathbb{R}^m$ que têm a forma $t_1 v_1 + \cdots + t_m v_m$, onde $t_1, \ldots, t_m \in [0, 1]$. Diremos também que $[v_1, \ldots, v_m]$ é o paralelepípedo que tem os vetores $v_1, \ldots, v_m$ como arestas. Seja $T\colon \mathbb{R}^m \to \mathbb{R}^m$ a transformação linear tal que $T \cdot e_1 = v_1, \ldots, T \cdot e_m = v_m$. O paralelepípedo $[v_1, \ldots, v_m]$ é a imagem por T do cubo unitário $C = [e_1, \ldots, e_m] = \{x \in \mathbb{R}^m; 0 \leq x_i \leq 1\}$. Por definição, temos vol. $C = 1$. Logo, pelo corolário acima, concluimos que vol. $[v_1, \ldots, v_m] = |\det.\ T| = |\det.\ (v_1, \ldots, v_m)|$, onde $(v_1, \ldots, v_m)$ é a matriz $m \times m$ cuja i-ésima coluna é o vetor v_i. Se um dos vetores v_i (digamos v_1, por simplicidade) é perpendicular aos demais v_j, então o volume do paralelepípedo é dado por vol. $[v_1, \ldots, v_m] = |v_1| \cdot \text{vol}_{m-1}[v_2, \ldots, v_m]$, onde vol_{m-1} significa o volume "$(m-1)$-dimensional" do paralelepípedo cujas arestas são $v_2, \ldots, v_m$. Isto é claro se $v_2, \ldots, v_m$ geram $\mathbb{R}^{m-1}$. O caso geral se reduz a este mediante uma transformação ortogonal (a qual preserva volumes de acordo com o Corolário). Evidentemente, se $v_1, \ldots, v_m$ são linearmente dependentes, (ou seja, se T não é invertível) então vol. $[v_1, \ldots, v_m] = |\det.\ T| = 0$.

Definiremos ainda o *volume orientado* do paralelepípedo $[v_1, \ldots, v_m] \subset \mathbb{R}^m$ pondo vol. $[v_1, \ldots, v_m] = \det.\ (v_1, \ldots, v_m)$. Neste caso, o volume possui sinal e a ordem em que são numeradas as arestas $v_1, \ldots, v_m$ é relevante. Os volumes orientados serão úteis no capítulo seguinte.

A demonstração do caso geral do teorema será precedida de dois lemas. O primeiro é de natureza topológica.

Lema 1. *Sejam X compacto, contido no aberto $U \subset \mathbb{R}^m$, e $\varphi\colon U \times U \to \mathbb{R}$ contínua, com $\varphi(x,x) = 1$ para todo $x \in X$. Dado $\varepsilon > 0$, pode-se obter $\delta > 0$ tal que $|\varphi(x,y) - 1| < \varepsilon$ quaisquer que sejam $x, y \in X$ com $|y - x| < \delta$.*

Demonstração: Se o lema fosse falso então, para cada $k \in N$, existiriam $x_k, y_k \in X$ tais que $|x_k - y_k| < 1/k$ e $|\varphi(x_k, y_k) - 1| \geq \varepsilon$. Passando a subseqüências, se necessário, (pois X é compacto) pode-se admitir que $\lim x_k = \lim y_k = a \in X$. Pela continuidade de φ, resultaria $|\varphi(a,a) - 1| \geq \varepsilon$, uma contradição.

Lema 2. *Sejam $U, V \subset \mathbb{R}^m$ abertos, $h\colon U \to V$ um difeomorfismo de classe C^1, $X \subset U$ um compacto J-mensurável e $N = N(h;X) = \sup.\{|h'(x)|; x \in X\}$. Então $h(X)$ é J-mensurável e* vol. $h(X) \leq N^m \cdot$ vol. X.

Demonstração: Consideremos em $\mathbb{R}^m$ a norma do máximo. Suponhamos inicialmente que $X = C$ seja um cubo de centro p, cuja aresta mede a, isto é, $C = \{x \in \mathbb{R}^m; |x - p| \leq a/2\}$. Pela desigualdade do Valor Médio, $x \in C \Rightarrow |h(x) - h(p)| \leq N \cdot a/2$. Isto significa que $h(C)$ está contido no cubo $D = \{y \in \mathbb{R}^m; |y - h(p)| \leq N \cdot a/2\}$, cuja aresta mede $N \cdot a$, logo vol. $h(C) \leq$ vol. $D = N^m \cdot a^m = N^m \cdot$ vol. C. No caso geral, tomemos arbitrariamente $\varepsilon > 0$. Sabemos que $|h'(x)|^m < N^m + \varepsilon$ para todo $x \in X$. Pela continuidade de h', existe um aberto U_0 tal que $X \subset U_0 \subset U$ e $|h'(y)|^m < N^m + \varepsilon$ para todo $y \in U_0$. Consideremos cubos $C_1, \ldots, C_k$ tais que $X \subset C_1 \cup \cdots \cup C_k \subset U_0$ e $\sum_{i=1}^{k}$ vol. $C_i <$ vol. $X + \varepsilon$. Para cada $i = 1, \ldots, k$, ponhamos $N_i = N(h; C_i) = \sup.\{|h'(y)|; y \in C_i\}$, o que dá $N_i^m \leq N^m + \varepsilon$. Temos $h(X) \subset \cup h(C_i)$, logo, como provamos acima, vol. $h(X) \leq \sum_{i=1}^{k}$ vol. $h(C_i) \leq \sum_{i=1}^{k} N_i^m \cdot$ vol. $C_i \leq (N^m + \varepsilon) \cdot \Sigma$ vol. $C_i \leq (N^m + \varepsilon)($vol. $X + \varepsilon)$. Como $\varepsilon > 0$ é arbitrário, concluimos que vol. $h(X) \leq N^m \cdot$ vol. X.

Teorema de Mudança de Variáveis. *Sejam $h\colon U \to V$ um difeomorfismo de classe C^1 entre abertos $U, V \subset \mathbb{R}^m$, $X \subset U$ um compacto J-mensurável e $f\colon h(X) \to \mathbb{R}$ uma função integrável. Então $f \circ h\colon X \to \mathbb{R}$ é integrável e*

$$\int_{h(X)} f(y)dy = \int_X f(h(x)) \cdot |\,\text{det. } h'(x)|dx.$$

Demonstração: Entre os conjuntos de pontos de descontinuidade de f e $f \circ h$ existe a relação $D_{f\circ h} = h^{-1}(D_f)$. Como D_f tem medida nula e h^{-1} é localmente Lipschitziana, $D_{f\circ h}$ tem medida nula, ou seja, $f \circ h$ é integrável. Para demonstrar a fórmula, notemos que basta considerar o caso em que $f \geq 0$, e que as decomposições pontilhadas de $h(X)$ têm a forma $\{h(X_1), \ldots, h(X_k); h(\xi_1), \ldots, h(\xi_k)\}$, onde $\{X_1, \ldots, X_k; \xi_1, \ldots, \xi_k\}$ é uma decomposição pontilhada de X. Portanto

$$\int_{h(X)} f(y)dy = \lim_{|D|\to 0} \sum_{i=1}^{k} f(h(\xi_i)) \cdot \text{vol. } h(X_i).$$

Dada a decomposição pontilhada D^* do conjunto X, ponhamos $T_i = h'(\xi_i)$, $N_i = \sup\,\{|T_i^{-1} \cdot h'(x)|; x \in X_i\}$ e $\eta_i = \eta_i(D) = N_i - 1$. Segue-se do Lema 1, com $\varphi(x,y) = |h'(y)^{-1} \cdot h'(x)|$, que para todo $\varepsilon > 0$ dado existe $\delta > 0$ tal que $|D| < \delta \Rightarrow |\eta_i| < \varepsilon$. Noutras palavras, $\lim_{|D|\to 0} \eta_i(D) = 0$. Usando o Caso Linear e o Lema 2, podemos escrever:

$$\begin{aligned}\text{vol. } h(X_i) &= \text{vol. } T_i \cdot T_i^{-1} \cdot h(X_i) = |\det.\ T_i| \cdot \text{vol. } T_i^{-1}\, h(X_i) \leq \\ &\leq |\det.\ T_i| N_i^m \cdot \text{vol. } X_i = |\det.\ h'(\xi_i)| \cdot (1+\eta_i)^m \cdot \text{vol. } X_i\,.\end{aligned}$$

Levando em conta que $f \geq 0$ e usando o Lema de Duhamel, obtemos

$$\begin{aligned}\int_{h(X)} f(y)dy &\leq \lim_{|D|\to 0} \Sigma\, f(h(\xi_i)) \cdot |\det.\ h'(\xi_i)| \cdot (1+\eta_i)^m \cdot \text{vol. } X_i = \\ &= \int_X f(h(x)) \cdot \det.\ |h'(x)| \cdot dx.\end{aligned}$$

Esta desigualdade contém a igualdade desejada. Com efeito, aplicando à última integral acima a mudança de variáveis $h^{-1} : h(X) \to X$ e observando que, para $x = h^{-1}(y)$, vale $|\det.\ h'(x)| \cdot |\det.\ (h^{-1})'(y)| = 1$, temos

$$\begin{aligned}&\int_X f(h(x)) \cdot |\det.\ h'(x)| dx \leq \\ &\leq \int_{h(X)} fh h^{-1}(y) \cdot |\det.\ h'(x)| \cdot |\det.\ (h^{-1})'(y)| dy = \\ &= \int_{h(X)} f(y)dy.\end{aligned}$$

APÊNDICE

Provaremos que toda transformação linear invertível $T\colon \mathbb{R}^m \to \mathbb{R}^m$ se escreve como produto $T = E_1 \cdot E_2 \dots E_k$ de transformações lineares *elementares* (necessariamente invertíveis) dos Tipos 1 e 2 definidos na demonstração do Caso Linear do teorema da mudança de variáveis.

Para começar, notamos que toda transformação linear da forma

$$x = (x_1, \dots, x_m) \mapsto E \cdot x = (x_1, \dots, x_{i-1}, \varphi(x), x_{i+1}, \dots, x_m),$$

onde $\varphi(x) = a_1x_1 + \cdots + a_mx_m$, é produto de uma transformação do Tipo 1 por uma do Tipo 2. Logo não há prejuízo em admitir as transformações da espécie acima entre as elementares. Observemos que det. $E = a_i$, logo E é invertível se, e somente se, $a_i \neq 0$.

Dada $T\colon \mathbb{R}^m \to \mathbb{R}^m$ invertível, mostraremos que é possível multiplicar T sucessivamente por transformações lineares elementares até obter a transformação identidade. Procedendo por indução, basta provar que se T é da forma

$$T \cdot x = (\varphi_1(x), \dots, \varphi_i(x), x_{i+1}, \dots, x_m),$$

onde $\varphi_r(x) = a_{r1}x_1 + \cdots + a_{rm}x_m$ $(r = 1, 2, \dots, i)$, então podemos encontrar transformações elementares E_1, E_2 tais que

$$TE_1E_2 \cdot x = (\xi_1(x), \dots, \xi_{i-1}(x), x_i, x_{i+1}, \dots, x_m).$$

Ora, as i primeiras linhas da matriz de T são os vetores $(a_{r1}, \dots, a_{rm})$ e as linhas seguintes são os vetores $e_{i+1}, \dots, e_m$ da base canônica de $\mathbb{R}^m$. Logo det. $T =$ det. (a_{rs}), $1 \leq r,\ s \leq i$. Como det. $T \neq 0$, deve existir algum s, com $1 \leq s \leq i$, tal que $a_{is} \neq 0$. Seja E_1 a transformação elementar (do Tipo 2) que consiste em permutar as coordenadas x_s e x_i. Considerando $S = T \cdot E_1$, temos

$$S \cdot x = (\psi_1(x), \dots, \psi_i(x), x_{i+1}, \dots, x_m),$$

com $\psi_i(x) = b_1x_1 + \cdots + b_mx_m$, onde $b_i \neq 0$. Portanto a transformação elementar $E \cdot x = (x_1, \dots, x_{i-1}, \psi_i(x), x_{i+1}, \dots, x_m)$ é invertível. A transformação $E_2 = E^{-1}$ é da mesma espécie que E e cumpre $\psi_i(E_2 \cdot x) = x_i$. Logo, pondo $\xi_j(x) = \psi_j(E_2 \cdot x)$, temos

$$TE_1E_2 \cdot x = SE_2 \cdot x = (\xi_1(x), \dots, \xi_{i-1}(x), x_i, x_{i+1}, \dots, x_m),$$

como queríamos demonstrar.

Exercícios

§1.

1.1. Sejam P e Q partições do bloco m-dimensional A. Um bloco m-dimensional B pertence à partição $P+Q$ se, e somente se, é interseção de um bloco de P com um bloco de Q.

1.2. Seja $f\colon A \to \mathbb{R}$ uma função definida no bloco m-dimensional A, tal que $m \leq f(x) \leq M$ para todo $x \in A$. Dada uma partição P de A, existe uma partição Q do bloco $(m+1)$-dimensional $A \times [m, M]$ tal que $S(f;P) - s(f;P)$ é a soma dos volumes dos blocos de Q que contêm pontos do gráfico de f.

1.3. Sejam $f, g\colon A \to \mathbb{R}$ limitadas no bloco A. Prove que

$$\underline{\int}_A f(x)dx + \underline{\int}_A g(x)dx \leq \underline{\int}_A [f(x)+g(x)]dx \leq$$
$$\leq \overline{\int}_A [f(x)+g(x)]dx \leq \overline{\int}_A f(x)dx + \overline{\int}_A g(x)dx.$$

Dê um exemplo em que todas as desigualdades acima são estritas. Prove também que $\underline{\int}_A c \cdot f(x)dx = c\underline{\int}_A f(x)dx$ se $c > 0$ e $\underline{\int}_A c \cdot f(x)dx = c\overline{\int}_A f(x)dx$ quando $c < 0$. Estabeleça o resultado análogo para a integral superior. Prove ainda que se $f(x) \leq g(x)$ para todo $x \in A$ então

$$\underline{\int}_A f(x)dx \leq \underline{\int}_A g(x)dx \quad \text{e} \quad \overline{\int}_A f(x)dx \leq \overline{\int}_A g(x)dx.$$

1.4. Sejam $f\colon A \to \mathbb{R}$, $g\colon B \to \mathbb{R}$ funções limitadas não-negativas nos blocos A, B. Defina $\varphi\colon A \times B \to \mathbb{R}$ pondo $\varphi(x,y) = f(x) \cdot g(y)$. Prove que

$$\overline{\int}_{A\times B} \varphi(z)dz = \overline{\int}_A f(x)dx \cdot \overline{\int}_B g(y)dy$$

e que vale um resultado análogo para integrais inferiores.

1.5. Se $f, g\colon A \to \mathbb{R}$ são integráveis, prove a *desigualdade de Schwarz*:

$$\left[\int_A f(x)g(x)dx\right]^2 \leq \int_A f(x)^2 dx \cdot \int_A g(x)^2 dx.$$

1.6. Dados os blocos m-dimensionais $B_1, \ldots, B_k$, contidos no bloco m-dimensional A, existe uma partição P de A tal que cada bloco B_i é reunião de blocos de P. Se cada interseção $B_i \cap B_j$ é (vazia ou) uma face comum a B_i e B_j, então P pode ser tomada tal que $B_i \in P$ para todo i.

1.7. Seja $f\colon A \to \mathbb{R}$ tal que $0 \leq f(x) \leq M$ para todo x pertencente ao bloco A. Considere o conjunto $C(f) = \{(x,y) \in A \times [0,M]; 0 \leq y \leq f(x)\}$. Mostre que

$$\overline{\int}_A f(x)dx = \overline{\int}_{A\times[0,M]} \chi_{C(f)}(x)dx.$$

Estabeleça o resultado análogo para integral inferior.

§2.

2.1. Se $X \subset \mathbb{R}^m$ tem medida nula então, para todo $Y \subset \mathbb{R}^n$, o produto cartesiano $X \times Y$ tem medida nula em $\mathbb{R}^{m+n}$.

2.2. Seja $f\colon [a,b] \to \mathbb{R}^n$ $(n > 1)$ um caminho contínuo retificável. Prove que $f([a,b])$ tem medida nula. [*Observação*: Existem "curvas de Peano". Cada uma delas é um caminho contínuo $f\colon [0,1] \to \mathbb{R}^n$, cuja imagem contém um cubo em $\mathbb{R}^n$.]

2.3. Dados $X \subset \mathbb{R}^p$ e $v \in \mathbb{R}^p$, seja $X + v = \{x + v; x \in X\}$. Sejam $M, N \subset \mathbb{R}^p$ superfícies de classe C^1 tais que dim. $M +$ dim. $N < p$. Mostre que é denso em $\mathbb{R}^p$ o conjunto dos vetores v, tais que $M + v$ e N são conjuntos disjuntos. [*Sugestão*: Prove que o conjunto dos pontos $x - y \in \mathbb{R}^p$, onde $x \in M$ e $y \in N$, tem medida nula.]

2.4. Diz-se que duas superfícies $M, N \subset \mathbb{R}^p$ se *intersectam transversalmente* em $\mathbb{R}^p$ quando, para todo $x \in M \cap N$, vale $\mathbb{R}^p = T_xM + T_xN$. [Note que a definição é fraseada de tal modo que $M \cap N = \varnothing \Rightarrow M$ e N se intersectam transversalmente.] Mostre que $M + v$ e N se intersectam transversalmente em $\mathbb{R}^p$ se, e somente se, $v \in \mathbb{R}^p$ é um valor regular da aplicação $\varphi\colon M \times N \to \mathbb{R}^p$, $\varphi(x,y) = y - x$. Conclua que, se dim. $M +$ dim. $N = p$ então o conjunto dos vetores $v \in \mathbb{R}^p$ tais que $M + v$ e N se intersectam transversalmente em $\mathbb{R}^p$ é denso em $\mathbb{R}^p$.

2.5. Para todo vetor unitário $u \in \mathbb{R}^p$, seja $E_u \subset \mathbb{R}^p$ o subespaço de dimensão $p-1$ formado pelos vetores perpendiculares a u e seja $\pi_u\colon \mathbb{R}^p \to E_u$ a projeção ortogonal cujo núcleo é formado pelos múltiplos de u. Dada uma superfície $M \subset \mathbb{R}^p$, de classe C^1, se $p > 2.\text{dim}.M + 1$ então é denso em S^{p-1} o conjunto dos vetores unitários $u \in \mathbb{R}^p$ tais que a projeção π_u aplica M injetivamente em E_u. [*Sugestão*: Considere $N = \{(x,y) \in M \times M; x \neq y\} \subset \mathbb{R}^{2p}$, defina $\varphi\colon N \to S^{p-1}$, $\varphi(x,y) = \dfrac{x-y}{|x-y|}$ e observe que $\pi_u|M$ é injetiva se, e somente se, u não pertence à imagem de φ.]

2.6. Seja $M \subset \mathbb{R}^p$ uma superfície de classe C^2. Se $2.\text{dim}.M < p$ então a reunião $X = \bigcup_{x \in M} T_xM$ dos espaços vetoriais tangentes a M em todos os seus pontos é um conjunto de medida nula em $\mathbb{R}^p$.

2.7. Seja $f\colon U \to \mathbb{R}$ de classe C^2 no aberto $U \subset \mathbb{R}^m$. Para cada $a \in \mathbb{R}^m$ considere a função $g = g_a\colon U \to \mathbb{R}$, definida por $g(x) = f(x) - \langle a, x\rangle$. Mostre que existe um conjunto $X \subset \mathbb{R}^m$ de medida nula tal que, para todo $a \in \mathbb{R}^m - X$, g é uma *função de Morse* (isto é, seus pontos críticos são todos não-degenerados). Conclua que, dada $f \in C^2$ e dado $\varepsilon > 0$, se U é limitado, existe uma função de Morse $g\colon U \to \mathbb{R}$ tal que $|f(x) - g(x)| < \varepsilon$ e $|df(x)\cdot v - dg(x)\cdot v| \leq \varepsilon |v|$ para quaisquer $x \in U$ e $v \in \mathbb{R}^m$.

2.8. Seja p um número real ≥ 0. Diz-se que um conjunto $X \subset \mathbb{R}^m$ tem *p-medida nula* (no sentido de Hausdorff) quando, para todo $\varepsilon > 0$ dado, pode-se escrever

$X = X_1 \cup \cdots \cup X_i \cup \ldots$ com $\sum\limits_{i=1}^{\infty}$ (diam. $X_i)^p < \varepsilon$. Escreve-se, então, $m_p(X) = 0$. Prove:

a) $X \subset \mathbb{R}^m$ tem m-medida nula se, e somente se, med. $X = 0$.

b) Um compacto $X \subset \mathbb{R}^m$ tem p-medida nula se, e somente se, para todo $\varepsilon > 0$, pode-se escrever $X = X_1 \cup \cdots \cup X_i$, com (diam. $X_1)^p + \cdots +$ (diam. $X_i)^p < \varepsilon$.

c) Se $M \subset \mathbb{R}^n$ é uma superfície de classe C^1 e dimensão m então $m_p(M) = 0$ se $p > m$ e $m_p(M) \neq 0$ se $p = m$.

2.9. Seja $K \subset \mathbb{R}^m \times [a,b]$ compacto. Para cada $t \in [a,b]$, seja $K_t = \{x \in \mathbb{R}^m; (x,t) \in K\}$. Se, para todo $t \in [a,b]$ tem-se med. $K_t = 0$ então med. $K = 0$ em $\mathbb{R}^{m+1}$.

§3.

3.1. Dê exemplo de uma função descontínua f para a qual a oscilação $\omega(f;x)$ é uma função contínua de x. Dê também exemplo de uma função f tal que $x \mapsto \omega(f;x)$ admita uma infinidade de pontos de descontinuidade.

3.2. Seja $f\colon X \to \mathbb{R}$ limitada no conjunto compacto $X \subset \mathbb{R}^m$. Mostre que existe um ponto $x_0 \in X$ no qual a oscilação da função f assume seu valor máximo. Dê um exemplo mostrando que o mínimo pode não ser atingido.

3.3. Indicando genericamente com a notação D_φ o conjunto dos pontos de descontinuidade de uma aplicação φ, mostre que se a composta $g \circ f$ faz sentido então $D_{g\circ f} \subset D_f \cup f^{-1}(D_g)$. Dê exemplo em que esta inclusão é estrita.

3.4. Se $f\colon A \to \mathbb{R}$, definida no bloco A, é integrável e $g\colon [a,b] \to \mathbb{R}$ é contínua num intervalo contendo $f(A)$ então $g \circ f\colon A \to \mathbb{R}$ é integrável. ("Uma função contínua de uma função integrável é integrável.")

3.5. Seja $f\colon A \to B$ contínua tal que $|f(x) - f(y)| \geq c|x - y|$ com $c > 0$ constante e $x, y \in A$ quaisquer. Prove que, para todo $g\colon B \to \mathbb{R}$ integrável, a composta $g \circ f\colon A \to \mathbb{R}$ é integrável.

3.6. Defina um homeomorfismo $f\colon [0,1] \to [0,1]$ que leve um conjunto de Cantor de medida positiva sobre o verdadeiro conjunto de Cantor C. Considere a função integrável $g = \chi_C\colon [0,1] \to \mathbb{R}$. Mostre que a composta $g \circ f\colon [0,1] \to \mathbb{R}$ não é integrável.

3.7. Se uma função $f\colon A \to \mathbb{R}$, limitada no retângulo $A \subset \mathbb{R}^m$, é integrável então, seu gráfico tem volume zero. E a recíproca?

3.8. Se $X \subset \mathbb{R}^m$ tem volume zero, o mesmo ocorre com $\overline{X}$. E medida nula?

3.9. Se uma seqüência de funções integráveis $f_k\colon A \to \mathbb{R}$ converge uniformemente para uma função $f\colon A \to \mathbb{R}$, então f é integrável e $\lim\limits_{k\to\infty} \int_A f_k(x)dx = \int_A f(x)dx$.

3.10. Seja $f\colon A \to \mathbb{R}$ limitada e ≥ 0 no bloco $A \subset \mathbb{R}^m$. A fim de que f seja integrável, é necessário e suficiente que o conjunto $C(f) = \{(x,y) \in \mathbb{R}^{m+1}; x \in A,\ 0 \leq y \leq f(x)\}$ seja J-mensurável. No caso afirmativo, tem-se $\int_A f(x)dx =$ vol. $C(f)$.

3.11. Sejam $f\colon X \to \mathbb{R}$ contínua, limitada no conjunto J-mensurável X e $x_0 \in X$. Para cada $n \in \mathbb{N}$, seja X_n um conjunto J-mensurável de volume positivo, tal que $X_n \subset X \cap B(x_0; 1/n)$. Prove que

$$\lim_{n\to\infty} \frac{1}{\text{vol. } X_n} \int_{X_n} f(x)dx = f(x_0).$$

3.12. Sejam $U \subset \mathbb{R}^m$ um aberto J-mensurável e (K_i) uma seqüência de compactos J-mensuráveis tais que $K_i \subset \text{ int } K_{i+1}$ para todo $i \in \mathbb{N}$ e $U = \cup K_i$. Prove que, para toda função integrável $f\colon U \to \mathbb{R}$, tem-se $\int_U f(x)dx = \lim\limits_{i\to\infty} \int_{K_i} f(x)dx$.

3.13. Se $f\colon A \to \mathbb{R}$ é integrável então, para todo $\delta > 0$, o conjunto $E_\delta = \{x \in A; \omega(f;x) \geq \delta\}$ tem volume zero.

3.14. Sejam $X, Y \subset \mathbb{R}^m$ conjuntos J-mensuráveis. Prove que o conjunto $X - Y = \{x \in X; x \notin Y\}$ é J-mensurável e que vol. $(X - Y) =$ vol. $X -$ vol. $(X \cap Y)$.

§4.

4.1. Seja $f\colon A \to \mathbb{R}$ uma função real arbitrária, definida num bloco A. Se existe o limite $\lim\limits_{|P|\to 0} \Sigma(f; P^*)$ então f é limitada.

4.2. Seja $f\colon X \to \mathbb{R}$ uma função integrável no conjunto J-mensurável X. Mostre que ainda se tem $\int_X f(x)dx = \lim\limits_{|D|\to 0} \Sigma(f; D^*)$, mesmo se na soma de Riemann $\Sigma(f; D^*)$ considerarmos apenas as parcelas que correspondem aos conjuntos X_i da decomposição D que não têm pontos em comum com a fronteira de X.

4.3. Sejam $f\colon A \to \mathbb{R}$ integrável no bloco A e $X \subset A$ um subconjunto J-mensurável. Para cada partição pontilhada P^* do bloco A, considere as seguintes somas de Riemann adaptadas ao conjunto X: $\Sigma_1(f; P^*) =$ soma das parcelas $f(\xi_j) \cdot$ vol. B_i tais que B_i está contido no conjunto X, $\Sigma_2(f; P^*) =$ soma das parcelas $f(\xi_i) \cdot$ vol. B_j tais que B_j tem algum ponto em comum com X. Prove que $\int_X f(x)dx = \lim\limits_{|D|\to 0} \Sigma_1(f; D^*) = \lim\limits_{|D|\to 0} \Sigma_2(f; D^*)$.

§5.

5.1. Sejam $\varphi\colon [a,b] \to \mathbb{R}$ e $\psi\colon [c,d] \to \mathbb{R}$ integráveis. A função $f\colon [a,b] \times [c,d] \to \mathbb{R}$, definida no retângulo $A = [a,b] \times [c,d]$ por $f(x,y) = \varphi(x)\psi(y)$, é integrável e $\int_A f(x,y)dxdy = \big(\int_a^b \varphi(x)dx\big)\big(\int_c^d \psi(y)dy\big)$.

5.2. Seja $T = \{(x,y,z) \in \mathbb{R}^3; x \geq 0, y \geq 0, z \geq 0, x + y + z \leq 1\}$. Mostre que vol. $T = \int_0^1 dx \int_0^{1-x} dz \int_0^{1-y-z} dy = 1/6$.

5.3. Seja $f\colon A \times B \to \mathbb{R}$ integrável no bloco $A \times B \subset \mathbb{R}^{m+n}$, produto cartesiano dos blocos $A \subset \mathbb{R}^m$ e $B \subset \mathbb{R}^n$. Prove que existe um conjunto de medida nula $X \subset A$ tal que, para todo $x \in A - X$, a função $f_x\colon B \to \mathbb{R}$, definida por $f_x(y) = f(x,y)$, é integrável.

5.4. Sejam $A = [0,1] \times [0,1]$, $A_0 = A - \{0\}$ e $f\colon A_0 \to \mathbb{R}$ a função contínua (ilimitada) definida por $f(x,y) = \dfrac{x-y}{(x+y)^3} \cdot$ Mostre que se tem

$$\int_0^1 dx \int_0^1 f(x,y)dy = \frac{1}{2} \quad \text{e} \quad \int_0^1 dy \int_0^1 f(x,y)dx = -\frac{1}{2}\,,$$

observando que, em ambos os casos, a segunda integração é imprópria.

5.5. Sejam $\varphi, \psi \colon [a,b] \to \mathbb{R}$ funções contínuas tais que $\varphi(x) \leq \psi(x)$ para todo $x \in [a,b]$. Mostre que o conjunto $X = \{(x,y) \in \mathbb{R}^2; x \in [a,b], \varphi(x) \leq y \leq \psi(x)\}$ é J-mensurável e que, para toda função contínua $f \colon X \to \mathbb{R}$, tem-se

$$\int_X f(x,y)dxdy = \int_a^b dx \int_{\varphi(x)}^{\psi(x)} f(x,y)dy.$$

5.6. (Princípio de Cavalieri.) Sejam $X, Y \subset \mathbb{R}^{m+1}$ conjuntos J-mensuráveis tais que, para cada $t \in \mathbb{R}$, as secções $X_t = \{x \in \mathbb{R}^m; (x,t) \in X\}$ e $Y_t = \{y \in \mathbb{R}^m; (y,t) \in Y\}$ são ainda J-mensuráveis e têm o mesmo volume em $\mathbb{R}^m$. Prove que vol. $X =$ vol. Y.

§6.

6.1. Seja $f \colon U \to \mathbb{R}^m$ de classe C^1 no aberto $U \subset \mathbb{R}^m$. Para algum $a \in U$, seja $f'(a) \colon \mathbb{R}^m \to \mathbb{R}^m$ um isomorfismo. Mostre que $\lim\limits_{r\to 0} \dfrac{\text{vol. } f(B[a;r])}{\text{vol. } B[a;r]} = |\det. \ f'(a)|$.

6.2. No exercício anterior, mostre que se $f'(a)$ não for um isomorfismo, então $\lim\limits_{r\to 0} \dfrac{\text{vol. } f(B[a;r])}{\text{vol. } B[a;r]} = 0$.

6.3. Seja $f \colon \mathbb{R}^m \to \mathbb{R}^m$ um difeomorfismo tal que $f(B) \subset B$, onde B é a bola unitária fechada de $\mathbb{R}^m$, e $|\det. \ f'(x)| < 1$ para todo $x \in B$. Prove que, para toda função contínua $g \colon B \to \mathbb{R}$ tem-se $\lim\limits_{n\to\infty} \int_{f^n(B)} g(x)dx = 0$, onde $f^n = f \circ f \circ \cdots \circ f$ (n fatores).

6.4. Seja $f \colon D \to \mathbb{R}$ uma função contínua no disco fechado $D \subset \mathbb{R}^2$, de raio 1 e centro na origem. Para todo $n \in \mathbb{N}$, mostre que $\int_D f = n^2 \cdot \int_D f(z^n)(x^2 + y^2)^{n-1}dxdy$ onde $z = x + iy$. Por outro lado, se $X \subset \mathbb{R}^2$ é um conjunto J-mensurável, contido num ângulo de vértice na origem medindo $2\pi/n$ radianos, e $g \colon X \to \mathbb{R}$ é uma função contínua limitada então

$$\int_Y g = n^2 \cdot \int_X g(z^n)(x^2 + y^2)^{n-1}, \text{ onde } Y = \{z^n; z \in X\}.$$

6.5. Seja $U \subset \mathbb{R}^m$ um aberto (limitado ou não). Uma *exaustão* de U é uma seqüência de compactos J-mensuráveis $K_i \subset U$ tais que $U = \cup K_i$ e $K_i \subset \text{int.}\, K_{i+1}$ para todo $i \in \mathbb{N}$. Dada uma função contínua $f \colon U \to \mathbb{R}$ (limitada ou não) diz-se que a integral imprópria $\int_U f(x)dx$ é *convergente* quando, para toda exaustão $U = \cup K_i$, existe o limite $\lim\limits_{i\to\infty} \int_{K_i} f(x)dx$ e seu valor é independente da exaustão tomada. Mostre que a integral imprópria $\int_U f(x)dx$ é convergente se, e somente se, $\int_U |f(x)|dx$ é também convergente.

6.6. Seja $f \colon U \to \mathbb{R}$ uma função contínua não-negativa no aberto $U \subset \mathbb{R}^m$. As seguintes afirmações são equivalentes:

1) Para alguma exaustão $U = \cup K_i$, existe o limite $\lim\limits_{i\to\infty} \int_{K_i} f(x)dx$;

2) Existe uma constante $c > 0$ tal que $\int_K f(x)dx \leq c$ para todo compacto $K \subset U$;

3) A integral imprópria $\int_U f(x)dx$ é convergente. [*Sugestão*: Use o Teorema de Borel-Lebesgue para mostrar que se $U = \cup K_i$ é uma exaustão e $K \subset U$ é compacto então existe $i \in \mathbb{N}$ tal que $K \subset K_i$.]

6.7. Seja $f\colon U \to \mathbb{R}$ definida em $U = \{(x, y) \in \mathbb{R}^2; x > 0, y > 0\}$ por $f(x, y) = e^{-(x^2+y^2)}$. Considere a exaustão $U = \cup L_k$, onde

$$L_k = \left\{ r \cdot e^{i\theta};\ \frac{1}{k} \le r \le k,\ \frac{1}{k} \le \theta \le \frac{\pi}{2} - \frac{1}{k} \right\},$$

use coordenadas polares e conclua que a integral imprópria $\int_U f(x, y)dxdy$ é convergente e vale $\pi/4$. Por outro lado, usando a exaustão $U = \cup K_i$, onde $K_i = \left[\frac{1}{i}, i\right] \times \left[\frac{1}{i}, i\right]$, mostre que $\int_U f(x, y)dxdy = \left(\int_0^\infty e^{-t^2}\, dt\right)^2$. Conclua que $\int_0^\infty e^{-t^2}\, dt = \frac{\sqrt{\pi}}{2}\cdot$

6.8. Seja $f\colon U \to \mathbb{R}$ definida em $U = \{(x, y) \in \mathbb{R}^2; x > 0, y > 0\}$ por $f(x, y) = \operatorname{sen}(x^2 + y^2)$. Use a fórmula $\operatorname{sen}(a + b) = \operatorname{sen} a \cdot \cos b + \cos a \cdot \operatorname{sen} b$ para mostrar que se $U = \cup K_i$ é uma exaustão de U por meio de quadrados com lados paralelos aos eixos então $\lim\limits_{i\to\infty} \displaystyle\int_{K_i} f(x)dx = \frac{\pi}{4}\cdot$ Por outro lado, usando a exaustão $U = \cup L_k$, onde

$$L_k = \left\{ r \cdot e^{i\theta} \in \mathbb{R}^2;\ \frac{1}{k} \le r \le k,\ \frac{1}{k} \le \theta \le \frac{\pi}{2} - \frac{1}{k} \right\},$$

mostre que não existe $\lim\limits_{k\to\infty} \int_{L_k} f(x)dx$.

6.9. Seja $f\colon U \times V \to \mathbb{R}$ contínua no produto dos abertos $U \subset \mathbb{R}^m$ e $V \subset \mathbb{R}^n$ (limitados ou não). Se a integral imprópria $\int_{U\times V} f$ é convergente então valem as igualdades

$$\int_U dx \int_V f(x, y)dy = \int_V dy \int_U f(x, y)dx = \int_{U\times V} f.$$

6.10. Use as coordenadas polares para mostrar que

$$\int_{\mathbb{R}^2} \frac{dxdy}{(x^2 + y^2 + 1)^{3/2}} = 2\pi.$$

[*Observação* (Para ser entendida depois da leitura do Capítulo VII.) Sejam $P, H \subset \mathbb{R}^3$, respectivamente, o plano $z = 1$ e o hemisfério superior aberto de centro 0 e raio 1. Considere o difeomorfismo $f\colon P \to H$, definido por $f(x, y, 1) = \dfrac{(x, y, 1)}{\sqrt{x^2 + y^2 + 1}}$ (restrição a P da projeção radial). O integrando acima é a restrição da forma elemento de ângulo sólido ao plano P. Logo é igual a $f^*\omega$, onde ω é a forma elemento de área de H. Portanto a integral acima representa a área de H, ou seja, é igual a 2π, resultado obtido sem cálculo algum.]

6.11. Sejam B_4 a bola fechada de raio 1 com centro na origem de $\mathbb{R}^4$ e B_3 a bola análoga em $\mathbb{R}^3$. Mostre que o volume de B_4 é igual a

$$2 \cdot \int_{B_3} \sqrt{1-(x^2+y^2+z^2)}dxdydz.$$

Tomando coordenadas esféricas $x = r\cos\varphi\cos\theta$, $y = r\cos\varphi \operatorname{sen}\theta$ e $z = r\operatorname{sen}\varphi$, mostre que o volume de B_4 é igual a

$$2\int_{-\pi/2}^{\pi/2} d\varphi \int_0^{2\pi} d\theta \int_0^1 r^2\cos\varphi\sqrt{1-r^2}dr$$

e conclua que vol. $B_4 = \pi^2/2$. Mais geralmente, mostre que o volume de uma bola de raio r em $\mathbb{R}^4$ é $\pi^2 r^4/2$.

6.12. Seja $V_m(r)$ o volume da bola de centro na origem e raio r em $\mathbb{R}^m$. Prove que se tem a relação indutiva $V_{n+1}(r) = 2\cdot V_n(1)\int_0^r (r^2-t^2)^{m/2}\,dt$ e conclua daí que $V_m(r) = r^m\pi^{m/2} \Big/ \left(\dfrac{m}{2}\right)!$ quando m é par e

$$V_m(r) = r^m\pi^{(m-1)/2}\cdot\frac{2^m\left(\frac{m-1}{2}\right)!}{m!}$$

se m é ímpar. Observe que destas fórmulas resulta um fato curioso:

$$\lim_{m\to\infty} V_m(r) = 0.$$

Capítulo VII

Integrais de Superfície

1 Formas alternadas

Neste parágrafo, desenvolveremos os fundamentos algébricos necessários para estudar as formas diferenciais de grau maior do que 1. As noções que vamos abordar pertencem à Álgebra Multilinear. Nosso maior interesse será focalizado no dual E^* de um espaço vetorial E, de dimensão finita sobre os reais.

Para o que se segue, convém deixar explícito que uma base num espaço vetorial E é uma *família* linearmente independente máxima $(e_i)_{i\in I}$, de elementos de E, com índices num conjunto I. Uma família com índices no conjunto $I = \{1, 2, \ldots, r\}$ é chamada uma *lista*, uma *seqüência* de r elementos, ou uma r-*lista*. Quando for $I = \{1, 2, \ldots, r\}$, escreveremos a base $(e_i)_{i\in I}$ como $\{e_1, \ldots, e_r\}$, que é uma notação tradicional, embora abusiva. A preocupação de considerar uma base como uma família, em vez de um conjunto, não é um simples pedantismo. Sem isto, por exemplo, não ficaria bem definida a matriz $a = (a_{ij})$ de uma transformação linear $A\colon E \to F$, relativamente às bases $(e_j)_{j\in J}$ em E e $(f_i)_{i\in I}$ em F. Como se sabe, para cada $i \in I$ e cada $j \in J$, o elemento a_{ij} (que se diz pertencer à i-ésima linha e à j-ésima coluna de a) é a i-ésima coordenada do vetor $A \cdot e_j$ relativamente à base (f_i). Ou seja, tem-se $A \cdot e_j = \sum\limits_{i\in I} a_{ij} f_i$ para cada $j \in J$. Se permutarmos os índices de uma das bases, obteremos outra matriz.

Dados os espaços vetoriais E, F, uma aplicação $\varphi\colon E \times \cdots \times E \to F$, definida no produto cartesiano de r fatores iguais a E, diz-se r-*linear*

quando seus valores $\varphi(v_1, \dots, v_r)$ dependem linearmente de cada uma das variáveis $v_1, \dots, v_r \in E$. Mais explicitamente, devemos ter

$$\varphi(v_1, \dots, v_i + w_i, \dots, v_r) = \varphi(v_1, \dots, v_i, \dots, v_r) + \\ + \varphi(v_1, \dots, w_i, \dots, v_r)$$

e

$$\varphi(v_1, \dots, \lambda \cdot v_i, \dots, v_r) = \lambda \cdot \varphi(v_1, \dots, v_i, \dots, v_r),$$

quaisquer que sejam $v_1, \dots, v_i,\ w_i, \dots, v_r \in E$ e $\lambda \in \mathbb{R}$.

Teorema 1. *Sejam $\varphi, \psi\colon E \times \cdots \times E \to F$ aplicações r-lineares e G um conjunto de geradores do espaço vetorial E. Se $\varphi(v_1,\dots,v_r) = \psi(v_1, \dots, v_r)$ qualquer que seja a r-lista $(v_1, \dots, v_r)$ de elementos de G, então $\varphi = \psi$.*

Demonstração: (Indução em r.) Se $r = 1$ então $\varphi, \psi\colon E \to F$ são aplicações lineares. Para todo $x \in E$, tem-se $x = \Sigma\, \alpha_i v_i$, $v_i \in G$. Portanto $\varphi(x) = \varphi(\Sigma\, \alpha_i v_i) = \Sigma\, \alpha_i \varphi(v_i) = \Sigma\, \alpha_i \psi(v_i) = \psi(\Sigma\, a_i v_i) = \psi(x)$. Supondo o teorema verdadeiro para $r-1$, introduzimos, para cada $v \in G$, as aplicações $(r-1)$-lineares φ_v, ψ_v, definidas por $\varphi_v(x_1, \dots, x_{r-1}) = \varphi(x_1, \dots, x_{r-1}, v)$ e $\psi_v(x_1, \dots, x_{r-1}) = \psi(x_1, \dots, x_{r-1}, v)$. Pela hipótese do teorema, φ_v e ψ_v assumem o mesmo valor em todas as listas de $r-1$ elementos de G. Por indução, concluimos que $\varphi_v = \psi_v$, isto é, que $\varphi(x_1, \dots, x_{r-1}, v) = \psi(x_1, \dots, x_{r-1}, v)$, quaisquer que sejam os elementos $x_1, \dots, x_{r-1} \in E$ e $v \in G$. Como todo elemento $x_r \in E$ é combinação linear de elementos de G, vemos que para $x_1, \dots, x_r \in E$ quaisquer vale

$$\varphi(x_1, \dots, x_r) = \varphi(x_1, \dots, x_{r-1}, \Sigma\, \alpha_i v_i) = \Sigma\, \alpha_i \cdot \varphi(x_1, \dots, x_{r-1}, v_i) = \\ = \Sigma\, \alpha_i \cdot \psi(x_1, \dots, x_{r-1}, v_i) = \psi(x_1, \dots, x_r),$$

isto é, $\varphi = \psi$.

Teorema 2. *Seja $\{e_1, \dots, e_m\}$ uma base do espaço vetorial E. Para cada r-lista $\sigma = (i_1, \dots, i_r)$ de inteiros entre 1 e m suponhamos dado um vetor $w_\sigma \in F$. Nestas condições, existe uma única aplicação r-linear $\varphi\colon E \times \cdots \times E \to F$ tal que $\varphi(e_{i_1}, e_{i_2}, \dots, e_{i_r}) = w_\sigma$, seja qual for a lista $\sigma = (i_1, \dots, i_r)$.*

Demonstração: Consideremos inicialmente $r = 2$. Para cada par ordenado (i, j) de inteiros, compreendidos entre 1 e m, é dado um vetor $w_{ij} \in F$. Quaisquer dois vetores $x, y \in E$ se escrevem como $x = \Sigma\, \alpha_i e_i$

e $y = \Sigma\, \beta_j e_j$, com os coeficientes α_i e β_j univocamente determinados. Definimos então $\varphi \colon E \times E \to F$ pondo $\varphi(x,y) = \Sigma\, \alpha_i \beta_j w_{ij}$. Evidentemente, φ é bilinear e cumpre $\varphi(e_i, e_j) = w_{ij}$. No caso geral, apenas a notação é mais elaborada. Dados arbitrariamente $x_1, \ldots, x_r \in E$, temos $x_i = \sum_{j=1}^{m} \alpha_{ij} e_j$ $(i = 1, \ldots, r)$, onde os α_{ij} são univocamente determinados. Pomos então:

$$\varphi(x_1, \ldots, x_r) = \sum_{\sigma} \alpha_{1i_1} \cdot \alpha_{2i_2} \ldots \alpha_{ri_r} \cdot w_\sigma \,,$$

a soma se estendendo a todas as r-listas $\sigma = (i_1, \ldots, i_r)$ de inteiros compreendidos entre 1 e m. A unicidade segue-se do Teorema 1.

Para definir uma aplicação r-linear basta, portanto, dar (arbitrariamente) seus valores em todas as r-listas de elementos de uma base prefixada. Neste caso, diz-se que a aplicação foi obtida estendendo-a linearmente a partir desses valores.

As operações usuais de soma de duas aplicações e produto de uma aplicação por um número fazem do conjunto das aplicações r-lineares $\varphi \colon E \times \cdots \times E \to F$ um espaço vetorial, que designaremos por $\mathcal{L}_r(E; R)$.

Uma aplicação r-linear $\varphi \colon E \times \cdots \times E \to F$ chama-se *alternada* quando se tem $\varphi(v_1, \ldots, v_r) = 0$ sempre que a seqüência $(v_1, \ldots, v_r)$ possuir repetições (isto é, existirem $i \neq j$ com $v_i = v_j$).

A fim de que $\varphi \in \mathcal{L}_r(E; F)$ seja alternada, é necessário e suficiente que seja *anti-simétrica*, isto é, que

$$\varphi(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_r) = -\varphi(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_r)$$

para quaisquer $v_1, \ldots, v_r \in E$.

Para provar isto escrevamos simplesmente o primeiro membro da igualdade acima como $\varphi(v_i, v_j)$. Então, φ alternada implica

$$\begin{aligned} 0 = \varphi(v_i + v_j, v_i + v_j) &= \varphi(v_i, v_i) + \varphi(v_i, v_j) + \varphi(v_j, v_i) + \\ &+ \varphi(v_j, v_j) = \varphi(v_i, v_j) + \varphi(v_j, v_i), \end{aligned}$$

logo $\varphi(v_i, v_j) = -\varphi(v_j, v_i)$, portanto φ é anti-simétrica.

Reciprocamente, se φ é anti-simétrica então $\varphi(v,v) = -\varphi(v,v)$, donde $2 \cdot \varphi(v,v) = 0$ e $\varphi(v,v) = 0$, logo φ é alternada.

Usaremos a notação $\mathcal{A}_r(E; F)$ para indicar o conjunto das aplicações r-lineares alternadas de E em F. Como se vê facilmente, $\mathcal{A}_r(E; F)$ é um subespaço vetorial de $\mathcal{L}_r(E; F)$.

Estaremos interessados, de modo especial, nas aplicações r-lineares alternadas $\varphi\colon E \times \cdots \times E \to \mathbb{R}$, que chamaremos *formas r-lineares alternadas* ou, simplesmente, *r-formas.* O espaço vetorial $\mathcal{A}_r(E;\mathbb{R})$ será indicado pela notação simplificada $\mathcal{A}_r(E)$.

Os elementos de $\mathcal{A}_r(E)$ são também chamados formas alternadas de *grau r*. Em particular, as formas alternadas de grau 1 sobre E são os funcionais lineares.

Para o que se segue, será conveniente convencionar que $\mathcal{A}_0(E) = \mathbb{R}$, isto é, que as *formas de grau zero* são os escalares.

Exemplo. 1) Toda aplicação linear $\varphi\colon E \to F$ é alternada, já que não é possível violar a condição de anti-simetria. Assim, $\mathcal{A}_1(E;F) = \mathcal{L}(E;F)$. Em particular, $\mathcal{A}_1(E) = E^* =$ dual de E.

2) Toda aplicação r-linear $\varphi\colon \mathbb{R}\times\cdots\times\mathbb{R}\to F$ é da forma $\varphi(t_1,\ldots,t_r) = t_1 t_2 \ldots t_i v$ onde $v = \varphi(1,\ldots,1)$. Quando $r > 1$, vemos que φ alternada $\Rightarrow v = 0 \Rightarrow \varphi = 0$. Assim, $\mathcal{A}_r(\mathbb{R};F) = \{0\}$ se $r > 1$.

3) Sejam $f, g \in E^*$. Definimos a forma bilinear alternada $\varphi\colon E\times E \to \mathbb{R}$ pondo $\varphi(u,v) = f(u)\cdot g(v) - f(v)\cdot g(u)$ para quaisquer $u, v \in E$. Evidentemente, se $g = c\cdot f$, a forma φ é identicamente nula. Mas, se f e g são linearmente independentes, podemos obter vetores $u, v \in E$ tais que $f(u) = 1,\ \ f(v) = 0,\ \ g(u) = 0,\ g(v) = 1$, donde $\varphi(u,v) = 1$ e $\varphi \neq 0$. Este primeiro exemplo não-trivial de forma alternada será generalizado logo mais abaixo, no Exemplo 6.

4) O *produto vetorial* $\varphi\colon \mathbb{R}^3\times\mathbb{R}^3 \to \mathbb{R}^3$, $\varphi(x,y) = x\times y$, é a aplicação bilinear definida de acordo com o Teorema 2, do seguinte modo. Tomamos a base canônica $\{e_1, e_2, e_3\} \subset \mathbb{R}^3$ e pomos $e_1 \times e_1 = e_2 \times e_2 = e_3 \times e_3 = 0$, $e_1 \times e_2 = -e_2 \times e_1 = e_3$, $e_2 \times e_3 = -e_3 \times e_2 = e_1$, $e_3 \times e_1 = -e_1 \times e_3 = e_2$. Dados arbitrariamente $x = (x_1, x_2, x_3)$ e $y = (y_1, y_2, y_3)$ temos

$$\begin{aligned}
x \times y &= (x_1e_1 + x_2e_2 + x_3e_3) \times (y_1e_1 + y_2e_2 + y_3e_3) = \\
&= (x_2y_3 - x_3y_2)e_1 + (x_3y_1 - x_1y_3)e_2 + (x_1y_2 - x_2y_1)e_3 = \\
&= (x_2y_3 - x_3y_2,\ x_3y_1 - x_1y_3,\ x_1y_2 - x_2y_1).
\end{aligned}$$

Vê-se imediatamente que $x \times x = 0$, logo o produto vetorial é uma aplicação bilinear alternada. Como se vê, os produtos internos $\langle x\times y, x\rangle$ e $\langle x \times y, y\rangle$ são nulos, isto é, o produto vetorial $x \times y$ é perpendicular ao plano gerado por x e y. (Se x e y forem linearmente dependentes, não gerarão um plano mas, neste caso, $x \times y = 0$.) Assim, por exemplo,

se $\varphi\colon U_0 \to U$ é uma parametrização numa superfície (bidimensional) $M \subset \mathbb{R}^3$ então, em cada ponto $(u, v) \in U_0$, o produto vetorial $\dfrac{\partial\varphi}{\partial u} \times \dfrac{\partial\varphi}{\partial v}$ é um vetor normal à superfície M no ponto $\varphi(u, v)$. Se escrevermos $\varphi(u, v) = (x(u, v), y(u, v), z(u, v))$ (olhe a mudança de significado para os símbolos x, y, z!) então o vetor normal é

$$\frac{\partial\varphi}{\partial u} \times \frac{\partial\varphi}{\partial v} = \left(\frac{\partial y}{\partial u}\frac{\partial z}{\partial v} - \frac{\partial z}{\partial u}\frac{\partial y}{\partial v}, \frac{\partial z}{\partial u}\frac{\partial x}{\partial v} - \frac{\partial x}{\partial u}\frac{\partial z}{\partial v}, \frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial y}{\partial u}\frac{\partial x}{\partial v} \right).$$

5) O *determinante* de uma matriz $m \times m$ poderá ser considerado como uma forma m-linear alternada $\det \in \mathcal{A}_m(\mathbb{R}^m)$ se, para $v_1, \ldots, v_m \in \mathbb{R}^m$ pusermos $\det(v_1, \ldots, v_m) =$ determinante da matriz $m \times m$ cujas colunas são os vetores v_i. Sendo o determinante de uma matriz igual ao de sua transposta, obteremos o mesmo resultado se tomarmos os vetores $v_1 \ldots v_m$ como linhas de uma matriz $m \times m$. Resultará do Teorema 5 abaixo que $\dim \mathcal{A}_m(\mathbb{R}^m) = 1$, logo toda forma m-linear alternada em $\mathbb{R}^m$ é um múltiplo constante do determinante.

6) A partir de r funcionais lineares $f_1, \ldots, f_r \in E^*$, obtemos a forma r-linear alternada $f_1 \wedge \cdots \wedge f_r\colon E \times \cdots \times E \to \mathbb{R}$, chamada o *produto exterior* desses funcionais, definida por

$$(f_1 \wedge \cdots \wedge f_r)(v_1, \ldots, v_r) = \det(f_i(v_j)),$$

onde à direita temos o determinante da matriz $r \times r$ cuja i-ésima linha é $(f_i(v_1), \ldots, f_i(v_r))$ e cuja j-ésima coluna é $(f_1(v_j), \ldots, f_r(v_j))$. [O fato de que $f_1 \wedge \cdots \wedge f_r \in \mathcal{A}_r(E)$ segue-se da linearidade de cada f_i e do exemplo acima.]

7) Dado o espaço vetorial E, seja $\Lambda\colon E^* \times \cdots \times E^* \to \mathcal{A}_r(E)$ a aplicação definida por $\Lambda(f_1, \ldots, f_r) = f_1 \wedge \cdots \wedge f_r$. Como o determinante de uma matriz $r \times r$ é uma forma r-linear alternada nos seus vetores-linha, vemos que Λ é uma aplicação r-linear alternada de E^* em $\mathcal{A}_r(E)$. As formas r-lineares alternadas que pertencem à imagem de Λ, ou seja, as do tipo $f_1 \wedge \cdots \wedge f_r$, onde $f_1, \ldots, f_r \in E^*$, chamam-se *decomponíveis.* Nem toda r-forma é decomponível, mas o Teorema 5 (adiante) mostra que todo elemento de $\mathcal{A}_r(E)$ pode ser escrito (de várias maneiras) como soma de formas decomponíveis. [Em geral, a imagem de uma aplicação r-linear não é um subespaço vetorial.]

Teorema 3. *Seja $\varphi\colon E \times \cdots \times E \to F$ uma aplicação r-linear alternada. Se $v_1, \ldots, v_r \in E$ são linearmente dependentes então $\varphi(v_1, \ldots, v_r) = 0$.*

Demonstração: Um dos vetores $v_1, \ldots, v_r$ é combinação linear dos demais, digamos $v_1 = \alpha_2 v_2 + \cdots + \alpha_r v_r$. Então

$$\varphi(v_1, \ldots, v_r) = \sum_{i=2}^{r} \alpha_i \cdot \varphi(v_i, v_2, \ldots v_r) = 0,$$

pois $\varphi(v_2, v_2, \ldots, v_r) = \varphi(v_3, v_2, v_3, \ldots, v_r) = \cdots = \varphi(v_r, v_2, \ldots, v_r) = 0$, já que φ é alternada.

Corolário 1. *O produto exterior $f_1 \wedge \cdots \wedge f_r$ é uma r-forma diferente de zero se, e somente se, $f_1, \ldots, f_r$ são linearmente independentes em E^*.*

Como a aplicação r-linear $\Lambda\colon (f_1, \ldots, f_r) \mapsto f_1 \wedge \cdots \wedge f_r$ é alternada, segue-se do Teorema que se $f_1 \wedge \cdots \wedge f_r \neq 0$ então $f_1, \ldots, f_r$ são linearmente independentes. Reciprocamente, se estes funcionais forem linearmente independentes, poderemos estendê-los a uma base de E^*. Seja $\{v_1, \ldots, v_m\} \subset E$ a base dual. Então, para i e j variando de 1 a r temos $f_i(v_j) = 0$ se $i \neq j$ e $f_i(v_i) = 1$. Logo $(f_i(v_j))$ é a matriz identidade $r \times r$ e daí segue que $(f_1 \wedge \cdots \wedge f_r)(v_1, \ldots, v_r) = 1$. Em particular, $f_1 \wedge \cdots \wedge f_r \neq 0$.

Corolário 2. *Se $r > \dim E$ então $\mathcal{A}_r(E) = \{0\}$.*

Com efeito, neste caso quaisquer r vetores em E são linearmente dependentes.

Teorema 4. *Sejam $\varphi, \psi\colon E \times \cdots \times E \to F$ aplicações r-lineares alternadas e $\{e_1, \ldots, e_m\}$ uma base de E. Se, para toda seqüência crescente $i_1 < \cdots < i_r$ de r inteiros compreendidos entre 1 e m, tivermos $\varphi(e_{i_1}, \ldots . e_{i_r}) = \psi(e_{i_1}, \ldots, e_{i_r})$ então $\varphi = \psi$.*

Demonstração: Seja $(j_1, \ldots, j_r)$ uma r-lista qualquer de inteiros entre 1 e m. Se houver elementos repetidos nessa lista, então

$$\varphi(e_{j_1}, \ldots, e_{j_r}) = \psi(e_{j_1}, \ldots, e_{j_r}) = 0$$

pois φ e ψ são alternadas. Se, porém, os termos da lista são todos distintos, então, por meio de sucessivas transposições, (trocas de posição entre 2 elementos apenas) podemos pôr os números $j_1, \ldots, j_r$ na ordem crescente $i_1 < \cdots < i_r$. Se são necessárias k transposições, a anti-simetria de φ e ψ, juntamente com a hipótese do teorema, nos dão:

$$\varphi(e_{j_1}, \ldots, e_{j_r}) = (-1)^k \varphi(e_{i_1}, \ldots, e_{i_r}) = (-1)^r \psi(e_{i_1}, \ldots, e_{i_r}) =$$
$$= \psi(e_{j_1}, \ldots, e_{j_r}).$$

Assim, as aplicações r-lineares φ e ψ cumprem a hipótese do Teorema 1, logo são iguais.

Seja $\{e_1, \ldots, e_m\}$ uma base do espaço vetorial E^*. Usaremos a notação $I = \{i_1 < \cdots < i_r\}$ para representar um subconjunto I, com r elementos, contido em $\{1, 2, \ldots, m\}$, cujos membros foram numerados na ordem crescente. Existem $\binom{m}{r} = \dfrac{m(m-1)\ldots(m-r+1)}{r!}$ desses conjuntos $I = \{i_1 < \cdots < i_r\}$. Para cada um deles, escreveremos

$$e_I = e_{i_1} \wedge \cdots \wedge e_{i_r}\,.$$

Seja $\{\bar{e}_1, \ldots, \bar{e}_m\} \subset E$ a base dual de $\{e_1, \ldots, e_m\} \subset E^*$. Dados $I = \{i_1 < \cdots < i_r\}$ e $J = \{j_1 < \cdots < j_r\}$ contidos em $\{1, 2, \ldots, m\}$, temos

$$e_I(\bar{e}_{j_1}, \ldots, \bar{e}_{j_r}) = \begin{cases} 0 & \text{se} \quad I \neq J \\ 1 & \text{se} \quad I = J \end{cases}$$

Com efeito, se $I \neq J$ então existe $i_k \in I - J$, logo $e_{i_k}(\bar{e}_j) = 0$ para todo $j \in J$ e daí $e_I(\bar{e}_{j_1}, \ldots, \bar{e}_{j_r}) = 0$ [determinante de uma matriz cuja k-ésima linha é nula]. Se, porém, $I = J$, então $e_I(\bar{e}_{j_1}, \ldots, \bar{e}_{j_r}) = 1$ [determinante da matriz identidade].

Mais geralmente, sejam $v_1, \ldots, v_r \in E$, com $v_j = \sum_{i=1}^{m} a_{ij}\bar{e}_i$, para $j = 1, \ldots, r$. Indiquemos com $a = (a_{ij})$ a matriz $m \times r$ cujas r colunas são as coordenadas de cada vetor v_j. Para cada subconjunto $I \subset \{1, 2, \ldots, m\}$, com r elementos, indiquemos com a_I a matriz $r \times r$ obtida de a escolhendo-se as r linhas cujos índices pertencem ao conjunto I. Para cada $i \in I$, temos $e_i(v_j) = a_{ij}$. Então

$$e_I(v_1, \ldots, v_r) = \det(e_{i_k}(v_j)) = \det(a_{i_k j}) = \det(a_I).$$

Teorema 5. *Seja $\{e_1, \ldots, e_m\}$ uma base de E^*. As r-formas $e_I = e_{i_1} \wedge \cdots \wedge e_{i_r}$, onde $I = \{i_1 < \cdots < i_r\}$ percorre os subconjuntos de $\{1, 2, \ldots, m\}$ com r elementos, constituem uma base de $\mathcal{A}_r(E)$. Em particular,* $\dim \mathcal{A}_r(E) = \binom{m}{r}$.

Demonstração: Dada arbitrariamente $\omega \in \mathcal{A}_r(E)$, ponhamos, para cada $I = \{i_1 < \cdots < i_r\}$, $\alpha_I = \omega(\bar{e}_{i_1}, \ldots, \bar{e}_{i_r})$, onde $\{\bar{e}_1, \ldots, \bar{e}_m\} \subset E$ é a base dual de $\{e_1, \ldots, e_m\}$. A r-forma $\varphi = \sum_I \alpha_I e_I$ é tal que, para toda seqüência crescente $j_1 < \cdots < j_r$ de inteiros compreendidos entre 1 e

m, tem-se $\varphi(\bar{e}_{j_1}, \ldots, \bar{e}_{j_r}) = \sum\limits_I \alpha_I e_I(\bar{e}_{j_1}, \ldots, \bar{e}_{j_r}) = \alpha_J = \omega(\bar{e}_{j_1}, \ldots, \bar{e}_{j_r})$. Segue-se do Teorema 4 que $\varphi = \omega$, ou seja, $\omega = \Sigma\, \alpha_I e_I$. Isto prova que as r-formas e_I geram $\mathcal{A}_r(E)$. Além disso, estas formas são linearmente independentes pois de uma combinação linear $\varphi = \Sigma\, \alpha_I e_I = 0$ tiramos, para todo $J = \{j_1 < \cdots < j_r\}$, $0 = \varphi(\bar{e}_{j_1}, \ldots, \bar{e}_{j_r}) = \alpha_J$.

O Teorema 5 é o fato mais importante a respeito de r-formas alternadas. Um caso particular importante ocorre quando $r = m$: se $m = \dim E$ então $\dim \mathcal{A}_m(E) = 1$. Isto quer dizer que, a menos de um fator constante, há apenas uma forma alternada de grau m sobre um espaço vetorial m-dimensional.

Resulta também do Teorema 5 que toda r-forma alternada se escreve (de modo não único) como soma de r-formas decompossíveis.

Tomemos uma base $\{e_1, \ldots, e_m\}$ em E^*, dual da base $\{\bar{e}_1, \ldots, \bar{e}_m\}$ em E. Suponhamos que os funcionais lineares $f_1, \ldots, f_r \in E^*$ se exprimam, em termos da base (e_j), como $f_i = \sum\limits_{j=1}^{m} a_{ij} e_j$ $(i = 1, \ldots, r)$.

Quais serão as coordenadas do produto exterior $f_1 \wedge \cdots \wedge f_r$ relativamente à base (e_j) do espaço $\mathcal{A}_r(E)$?

Sabemos que existe uma expressão única $f_1 \wedge \cdots \wedge f_r = \sum\limits_J \alpha_J e_J$, com $\alpha_J = (f_1 \wedge \cdots \wedge f_r)(\bar{e}_{j_1}, \ldots, \bar{e}_{j_r}) = \det(f_i(\bar{e}_{j_k}))$ para todo $J = \{j_1 < \cdots < j_r\} \subset \{1, \ldots, m\}$.

Ora, a matriz $a = (a_{ij})$, com r linhas e m colunas, determinada pelas coordenadas dos funcionais f_i em relação à base (e_j), é caracterizada pelas relações $a_{ij} = f_i(\bar{e}_j)$, onde $1 \leq i \leq r$ e $1 \leq j \leq m$. Para todo subconjunto $J \subset \{1, 2, \ldots, m\}$, com r elementos, essa matriz possui uma submatriz $r \times r$, que representaremos por a_J , a qual é formada selecionando-se as r colunas da matriz $a = (a_{ij})$ cujos índices j pertencem ao conjunto J. Tem-se então $\alpha_J = \det(a_J)$. Assim:

$$f_i = \sum_{j=1}^{m} a_{ij} e_j \Rightarrow f_1 \wedge \cdots \wedge f_r = \sum_J \det(a_J) e_J\,,$$

a soma estendendo-se a todos os subconjuntos $J \subset \{1, \ldots, m\}$ com r elementos.

Em particular, se $r = m$, obtemos

$$f_1 \wedge \cdots \wedge f_m = \det(a) \cdot e_1 \wedge \cdots \wedge e_m\,,$$

onde $a = (a_{ij})$ é a matriz de passagem de $\{e_1, \ldots, e_m\}$ para $\{f_1, \ldots, f_m\}$.

Usaremos agora este resultado para estabelecer uma identidade útil sobre determinantes.

Exemplo 8 (*Identidade de Lagrange*). Sejam $a = (a_{ij})$ uma matriz $m \times r$, com $m \geq r$ e a^* sua transposta. Então

$$\det(a^* \cdot a)) = \sum_J [\det(a_J)]^2,$$

onde J percorre todos os subconjuntos de $\{1, \ldots, m\}$ com r elementos e a_J é a matriz $r \times r$ obtida de a escolhendo-se as r linhas cujos índices pertencem a J.

Com efeito, sejam $v_1, \ldots, v_r \in \mathbb{R}^m$ os vetores-coluna da matriz a.

Para cada $j = 1, \ldots, r$, temos $v_j = (a_{1j}, a_{2j}, \ldots, a_{mj})$. Se $\{e_1, \ldots, e_m\} \subset (\mathbb{R}^m)^*$ é a base canônica, temos $e_k(v_j) = a_{kj}$ para $1 \leq k \leq m$ e $1 \leq j \leq r$ quaisquer. Como vimos acima, para todo subconjunto $J \subset \{1, 2, \ldots, m\}$ com r elementos, temos $e_J(v_1, \ldots, v_r) = \det(a_J)$. Consideremos, em seguida, os funcionais $f_1, \ldots, f_r \in (\mathbb{R}^m)^*$, dados por $f_i = \sum\limits_{k=1}^{m} a_{ki}e_k$. Então, para $1 \leq i, j \leq r$, quaisquer:

$$f_i(v_j) = \sum_k a_{ki}e_k(v_j) = \sum_k a_{ki}a_{kj}\,,$$

donde $f_i(v_j)$ é o (i, j)-ésimo elemento da matriz produto $a^* \cdot a$. Segue-se daí, e da expressão $f_1 \wedge \cdots \wedge f_r = \sum\limits_J \det(a_J)e_J$, que

$$\det(a^*a) = \det(f_i(v_j)) = (f_1 \wedge \cdots \wedge f_r)(v_1, \ldots, v_r) =$$
$$= \sum_J \det(a_J)e_J(v_1, \ldots, v_r) = \sum_J [\det(a_J)]^2.$$

Toda transformação linear $A\colon E \to F$ possui uma transposta $A^*\colon F^* \to E^*$, definida por $(A^* \cdot \omega)(v) = \omega(A \cdot v)$ para todo $\omega \in F^*$ e todo $v \in E$. Esta noção se generaliza. Para todo r, a transformação linear $A\colon E \to F$ determina uma nova transformação linear $A^*\colon \mathcal{A}_r(F) \to \mathcal{A}_r(E)$, definida por

$$(A^* \cdot \omega)(v_1, \ldots, v_r) = \omega(A \cdot v_1, \ldots, A \cdot v_r),$$

para quaisquer $\omega \in \mathcal{A}_r(F)$ e $v_1, \ldots, v_r \in E$. A transformação A^* diz-se *induzida* por A nas formas de grau r. A forma $A^* \cdot \omega$ chama-se

também a *forma induzida* por A em E ou o *pull-back* da forma ω para o espaço E.

Sejam $\{e_1, \dots, e_m\} \subset E^*$ e $\{f_1, \dots, f_n\} \subset F^*$ bases, duais respectivamente das bases $\{\bar{e}_1, \dots, \bar{e}_m\} \subset E$ e $\{\bar{f}_1, \dots, \bar{f}_n\} \subset F$. Seja $a = (a_{ij})$ a matriz (com n linhas e m colunas) da transformação linear $A\colon E \to F$ em relação a essas bases, isto é, $A \cdot \bar{e}_j = \sum_{i=1}^{n} a_{ij}\bar{f}_i$, para $j = 1, \dots, m$. Sabemos que a matriz de $A^*\colon F^* \to E^*$ é a transposta da matriz de A, pois $A^* \cdot f_i = \sum_{j=1}^{m} a_{ij}e_j$ $(i = 1, \dots, n)$. Determinaremos agora a matriz da transformação linear induzida $A^*\colon \mathcal{A}_r(F) \to \mathcal{A}_r(E)$ relativamente às bases (f_I) e (e_J), determinadas de acordo com o Teorema 5.

Temos $A^* \cdot f_I = \sum_J \alpha_{IJ}e_J$, onde

$$\begin{aligned} \alpha_{IJ} = (A^* \cdot f_I)(\bar{e}_{j_1}, \dots, \bar{e}_{j_r}) = f_I \cdot (A \cdot e_{j_1}, \dot{;} A \cdot e_{j_r}) = \\ = \det(f_{i_\lambda}(A \cdot \bar{e}_{j_\mu})) = \det(a_{i_\lambda j_\mu}) \end{aligned}$$

se $I = \{i_1 < \cdots < i_r\}$ e $J = \{j_1 < \cdots < j_r\}$. Indicando com a_{IJ} a submatriz $r \times r$ que consiste em selecionar da matriz a cada elemento a_{ij} cujo índice i pertence a I e cujo índice j pertence a J, temos $\alpha_{IJ} = \det(a_{IJ})$.

A matriz de $A^*\colon \mathcal{A}_r(F) \to \mathcal{A}_r(E)$ possui $\binom{m}{r}$ linhas e $\binom{n}{r}$ colunas. Ela é transposta da matriz cujos elementos $\alpha_{IJ} = \det(a_{IJ})$ têm como índices de linha os subconjuntos $I = \{i_1 < \cdots < i_r\} \subset \{1, 2, \dots, n\}$ e como índices de coluna os subconjuntos $J = \{j_1 < \cdots < j_r\} \subset \{1, 2, \dots, m\}$.

Em particular, se $r = m = n$, então a transformação linear $A^*\colon \mathcal{A}_m(F) \to \mathcal{A}_m(E)$ é tal que $A^* \cdot (f_1 \wedge \cdots \wedge f_m) = \det(a) \cdot (e_1 \wedge \cdots \wedge e_m)$, onde $a = (a_{ij})$ é a matriz de $A\colon E \to F$ acima considerada.

Mais particularmente ainda, dada uma transformação linear $A\colon E \to E$ do espaço vetorial m-dimensional E em si mesmo, então det. A tem sentido (independentemente da escolha de bases) e a transformação linear induzida $A^*\colon \mathcal{A}_m(E) \to \mathcal{A}_m(E)$ é simplesmente $A^* \cdot \omega = (\det.\ A) \cdot \omega$.

Vejamos agora como mudam as coordenadas de uma forma $\omega \in \mathcal{A}_r(E)$ quando se faz uma mudança de bases em E.

Sejam $\{\bar{e}_1, \dots, \bar{e}_m\}$ e $\{\bar{f}_1, \dots, \bar{f}_m\}$ bases em E, relacionadas pelas equações $\bar{e}_j = \sum_{i=1}^{m} a_{ij}\bar{f}_i$ $(j = 1, \dots, m)$. Suas bases duais, $\{e_1, \dots, e_m\}$ e

$\{f_1, \ldots, f_m\}$, em E^*, cumprem as relações $f_i = \sum_{j=1}^{m} a_{ij} e_j$ $(i = 1, \ldots, m)$. Para quaisquer subconjuntos $I, J \subset \{1, \ldots, m\}$, com r elementos, indicamos com a_{IJ} a submatriz $r \times r$ da matriz $a = (a_{ij})$, formada pelos elementos a_{ij} com $i \in I$ e $j \in J$. Pelo visto acima, temos $f_I = \sum_J \det(a_{IJ}) e_J$.

Assim, se uma forma $\omega \in \mathcal{A}_r(E)$ admite as expressões $\omega = \sum_J \alpha_J e_J = \sum_I \beta_I f_I$ relativametne às bases (e_J) e (f_I), temos

$$\omega = \sum_J \alpha_J e_J = \sum_I \sum_J \beta_I \det(a_{IJ}) e_J = \sum_J \left[\sum_I \det(a_{IJ}) \beta_I \right] e_J .$$

Comparando os coeficientes de e_J obtemos

$$\alpha_J = \sum_I \det(a_{IJ}) \beta_I .$$

Nos termos do Cálculo Tensorial clássico, uma r-forma alternada seria dada, em cada base de E, por meio de suas coordenadas α_J , de tal modo que uma mudança de base em E determinaria uma mudança de coordenadas para a forma, de acordo com a expressão acima.

Convém notar o caso particular em que $\omega \in \mathcal{A}_m(E)$, $m = \dim E$. Temos então as bases $\{e\}$ e $\{f\}$ em $\mathcal{A}_m(E)$, com $e = e_1 \wedge \cdots \wedge e_m$, $f = f_1 \wedge \cdots \wedge f_m$ e $f = (\text{det. } a) \cdot e$, onde $a = (a_{ij})$ é a matriz de passagem. Uma m-forma arbitrária $\omega \in \mathcal{A}_m(E)$ se escreve como $\omega = \alpha \cdot e = \beta \cdot f$, com $\alpha = (\text{det. } a) \cdot \beta$, lembrando-se sempre as relações $f_i = \sum_{j=1}^{m} a_{ij} e_j$, que definem a matriz a.

Exemplo 9. *A forma elemento de volume.* Seja E um espaço vetorial de dimensão m, orientado e munido de um produto interno. [Orientar um espaço vetorial é escolher nele uma base, chamá-la de "positiva"e declarar também positivas todas as demais bases que dela se obtenham por meio de uma matriz de passagem com determinante > 0. As outras bases serão chamadas negativas.] Definiremos uma forma $\omega \in \mathcal{A}_m(E)$, chamada o *elemento de volume de E*. Primeiramente tomamos uma base ortonormal positiva $\{e_1, \ldots, e_m\}$ em E. Dada a seqüência de vetores $v_1, \ldots, v_m \in E$, tem-se $v_j = \sum_{i=1}^{m} a_{ij} e_i$, para cada $j = 1, \ldots, m$. Seja $a = (a_{ij})$ a matriz $m \times m$ assim obtida. Por definição, pomos

$$\omega(v_1, \ldots, v_m) = \text{det. } a.$$

Evidentemente a igualdade acima define uma forma $\omega \in \mathcal{A}_m(E)$. Resta provar que ω independe da escolha de base que fizemos. Para isto, introduzimos a *matriz de Gram* $g = (\langle v_i, v_j \rangle)$, na qual o elemento situado na i-ésima linha e na j-ésima coluna é o produto interno $\langle v_i, v_j \rangle$. Como

$$\langle v_i, v_j \rangle = \langle \sum_k a_{ki} e_k \,, \sum_s a_{sj} e_s \rangle = \sum_k a_{ki} a_{kj} \,,$$

vemos que $g = a^* \cdot a$, onde a^* é a transposta da matriz a. Daí resulta que det. $g = (\text{det. } a)^2$. Em particular, $\det \cdot g \geq 0$, sendo det. $g = 0 \Leftrightarrow$ os vetores $v_1, \ldots, v_m$ são linearmente dependentes. Concluimos que

$$\omega(v_1, \ldots, v_m) = \pm\sqrt{\det(\langle v_i, v_j \rangle)},$$

onde o sinal $+$ ou $-$ é o sinal de det. a. Assim, $\omega(v_1, \ldots, v_m) > 0$ quando os vetores $v_1, \ldots, v_m$ formam (nesta ordem) uma base positiva de E e $\omega(v_1, \ldots, v_m) < 0$ se a base $v_1, \ldots, v_m$ é negativa. (No caso de $v_1, \ldots, v_m$ serem linearmente dependentes, tem-se $\omega(v_1, \ldots, v_m) = 0$.)

A igualdade $\omega(v_1, \ldots, v_m) = \pm\sqrt{\text{det. } g}$ mostra que a definição de ω independe da escolha de uma base.

No caso $E = \mathbb{R}^m$, $|\text{det. } a|$ é o volume do paralelepípedo que tem como arestas os vetores $v_1, \ldots, v_m$, de modo que $\omega(v_1, \ldots, v_m)$ é o "volume orientado" desse paralelepípedo, ou seja, um volume dotado de sinal. No caso geral, tomaremos $\omega(v_1, \ldots, v_m)$ como definição desse volume.

Exemplo 10. Generalizando o Exemplo 4, definiremos o *produto vetorial* $= v_1 \times \cdots \times v_m$ de m vetores em $\mathbb{R}^{m+1}$ como o vetor $v \in \mathbb{R}^{m+1}$ tal que, para todo $x \in \mathbb{R}^{m+1}$, $\langle x, v \rangle = \det(x, v_1, \ldots, v_m)$ = determinante da matriz $(m+1) \times (m+1)$ cujas colunas são os vetores $x, v_1, \ldots, v_m$, nesta ordem. As coordenadas dos vetores $v_j = \sum_{i=1}^{m+1} a_{ij} e_i$ $(j = 1, \ldots, m)$ relativamente à base canônica de $\mathbb{R}^{m+1}$ formam uma matriz $a = (a_{ij})$ com $m+1$ linhas e m colunas. Indicaremos com $a_{(i)}$ a matriz $m \times m$ obtida de a por omissão de sua i-ésima linha. Resulta da definição do produto vetorial que, para todo $i = 1, \ldots, m+1$:

$$\langle e_i, v_1 \times \cdots \times v_m \rangle = \det(e_i, , v_1, \ldots, v_m) = (-1)^{i+1} \det \cdot a_{(i)} \,,$$

onde a última igualdade decorre da expansão de $\det(e_i, v_1, \ldots, v_m)$ rela-

tivamente à sua primeira coluna. Portanto

$$v_1 \times \cdots \times v_m = \sum_{i=1}^{m+1} (-1)^{i+1} \text{ det. } a_{(i)} \cdot e_i$$

é a expressão do vetor $v_1 \times \cdots \times v_m$ em termos da base canônica de $\mathbb{R}^{m+1}$. Daí, ou diretamente da definição, vemos que $(v_1,\ldots,v_m) \mapsto v_1 \times \cdots \times v_m$ é uma aplicação m-linear alternada de $\mathbb{R}^{m+1} \times \cdots \times \mathbb{R}^{m+1}$ em $\mathbb{R}^{m+1}$. O produto vetorial é caracterizado pelas seguintes propriedades:

0) $v_1 \times \cdots \times v_m = 0$ se os vetores $v_1, \ldots, v_m$ forem linearmente dependentes;

1) $v_1 \times \cdots \times v_m$ é perpendicular a cada v_j ;

2) $|v_1 \times \cdots \times v_m|$ é o volume (m-dimensional) do paralelepípedo gerado pelos vetores $v_1, \ldots, v_m$; em particular, $v_1 \times \cdots \times v_m \neq 0$ se $v_1, \ldots, v_m$ forem linearmente independentes;

3) det. $(v_1 \times \cdots \times v_m, v_1, \ldots, v_m) > 0$ se $v_1, \ldots, v_m$ forem linearmente independentes.

Constatemos primeiro estes fatos: 0) decorre de $v_1 \times \cdots \times v_m$ ser m-linear alternada como função dos v_i ; 1) segue-se da definição, pois $\langle v_i, v_1 \times \cdots \times v_m \rangle = \text{det.}\,(v_i, v_1, \ldots, v_m) = 0$ por ser um determinante com duas colunas iguais; quanto a 2), sabemos que o volume do paralelepípedo $[v_1, \ldots, v_m]$ é igual a $\sqrt{\text{det. } g}$, onde a matriz de Gram g tem como (i,j)-ésimo elemento o produto interno $\langle v_i, v_j \rangle = \sum_{k=1}^{m+1} a_{ki} a_{kj}$, portanto $g = a^* \cdot a$. Assim, pela identidade de Lagrange, podemos escrever:

$$\begin{aligned} \text{vol.}\,[v_1, \ldots, v_m] &= \sqrt{\text{det. } g} = \sqrt{\text{det.}\,(a^* a)} = \\ &= \sqrt{\Sigma\,\text{det.}\,(a_{(i)})^2} = |v_1 \times \cdots \times v_m|. \end{aligned}$$

Finalmente, para provar 3), basta observar que, por definição, det. $(v_1 \times \cdots \times v_m, v_1, \ldots, v_m) = |v_1 \times \cdots \times v_m|^2$.

Em seguida, mostremos que as propriedades acima de fato caracterizam o produto vetorial. Por 0), basta considerar o caso em que $v_1, \ldots, v_m$ são linearmente independentes. A propriedade 1) determina a direção do vetor $v_1 \times \cdots \times v_m$, isto é, a reta que passa pela origem e o contém. A propriedade 2) dá seu comprimento, enquanto 3) diz qual dos dois vetores não-nulos com aquele comprimento, sobre a reta dada, é o produto vetorial: ele é tal que $(v_1 \times \cdots \times v_m, v_1, \ldots, v_m)$ é uma base positiva de $\mathbb{R}^{m+1}$.

O produto vetorial $v_1 \times \cdots \times v_m$ pode ainda ser definido quando $v_1, \ldots, v_m$ são m vetores num espaço vetorial E, orientado e dotado de um produto interno. Basta tomar uma base ortonormal positiva $\{e_1, \ldots, e_{m+1}\}$ em E, escrever $v_j = \sum_i a_{ij} e_i$ para cada $j = 1, \ldots, m$ e pôr $v_1 \times \cdots \times v_m = \Sigma(-1)^{i+1}$ det. $a_{(i)} \cdot e_i$. Verifica-se, como antes, que valem as propriedades 0), 1), 2) e 3) acima, logo $v_1 \times \cdots \times v_m$ não depende da base ortonormal positiva escolhida para definí-lo. O produto vetorial é uma aplicação m-linear alternada $E \times \cdots \times E \to E$.

A seguir, definiremos o *produto exterior de uma r-forma por uma s-forma*, dando como resultado uma $(r+s)$-forma. Mais precisamente, queremos obter uma aplicação bilinear

$$\varphi \colon \mathcal{A}_r(E) \times \mathcal{A}_s(E) \to \mathcal{A}_{r+s}(E)$$

tal que $\varphi(f_1 \wedge \cdots \wedge f_r, g_1 \wedge \cdots \wedge g_s) = f_1 \wedge \cdots \wedge f_r \wedge g_1 \wedge g_1 \wedge \cdots \wedge g_s$, sejam quais forem os funcionais lineares $f_1, \ldots, f_r,\ g_1, \ldots, g_s \in E^*$. Como $\mathcal{A}_r(E)$ e $\mathcal{A}_s(E)$ são gerados por formas decomponíveis, se existir uma aplicação bilinear φ com a propriedade acima, ela será única. Podemos então usar escolhas arbitrárias para definir φ. Se a relação acima for válida, as escolhas que fizermos serão irrelevantes.

Tomemos uma base $\{e_1, \ldots, e_m\}$ em E^*. Para $I = \{i_1 < \cdots < i_r\}$ e $J = \{j_1 < \cdots < j_s\}$ contidos em $\{1, \ldots, m\}$, ponhamos

$$\varphi(e_I, e_J) = e_{i_1} \wedge \cdots \wedge e_{i_r} \wedge e_{j_1} \wedge \cdots \wedge e_{j_s}$$

e estendamos φ a uma aplicação bilinear de $\mathcal{A}_r(E) \times \mathcal{A}_s(E)$ em $\mathcal{A}_{r+1}(E)$, de acordo com o Teorema 2. Consideremos o diagrama

onde

$$\Lambda(f_1 \wedge \cdots \wedge f_r, g_1 \wedge \cdots \wedge g_s) = f_1 \wedge \cdots \wedge f_r \wedge g_1 \wedge \cdots \wedge g_s$$

$$\text{e} \qquad \alpha(f_1, \ldots, f_r, g_1, \ldots, g_s) = (f_1 \wedge \cdots \wedge f_r, g_1 \wedge \cdots \wedge g_s).$$

Devemos provar que $\varphi \circ \alpha = \Lambda$. Pelo Teorema 1, basta verificar que

$$(*) \qquad \varphi(e_{i_1} \wedge \cdots \wedge e_{i_r}, e_{j_1} \wedge \cdots \wedge e_{j_s}) = \\ = e_{i_1} \wedge \cdots \wedge e_{i_r} \wedge e_{j_1} \wedge \cdots \wedge e_{j_s}$$

para *qualquer* $(r+s)$-lista $(e_{i_1}, \ldots, e_{i_r}, e_{j_1}, \ldots, e_{j_s})$ de elementos da base escolhida.

A igualdade (*) é evidente quando uma das seqüências $(i_1, \ldots, i_r)$ ou $(j_1, \ldots, j_s)$ possui repetições, pois ambos os membros são então iguais a zero. Se nenhuma das seqüências tem repetição e, além disso, vale $i_1 < \cdots < i_r$ e $j_1 < \cdots < j_s$, então a igualdade (*) é a própria definição de φ. Finalmente, se as seqüências não possuem repetição mas pelo menos uma delas não está em ordem crescente, podemos levar ambas à ordem crescente mediante sucessivas transposições. Como cada transposição entre dois índices i ou entre dois índices j muda o sinal de ambos os membros de (*), concluimos que a igualdade em questão é válida em qualquer caso.

Dadas $\omega \in \mathcal{A}_r(E)$ e $\bar{\omega} \in \mathcal{A}_s(E)$, escreveremos $\omega \wedge \bar{\omega} \in A_{fr+s}(E)$ em vez de $\varphi(\omega, \bar{\omega})$. O produto exterior $\omega \wedge \bar{\omega}$, além de bilinear, goza ainda da propriedade de anti-comutatividade abaixo:

$$\bar{\omega} \wedge \omega = (-1)^{rs}\, \omega \wedge \bar{\omega} \quad \text{se} \quad \omega \in A_r(E) \text{ e } \bar{\omega} \in \mathcal{A}_s(E).$$

Isto é óbvio se ω e $\bar{\omega}$ são decomponíveis e o caso geral se reduz a este, pois toda forma alternada é soma de formas decomponíveis.

O produto exterior é, ainda, associativo. A relação $\alpha \wedge (\beta \wedge \gamma) = (\alpha \wedge \beta) \wedge \gamma$ é evidente para formas decomponíveis, logo é válida em geral, por distributividade.

Devemos ainda observar que se $A \colon E \to F$ é uma transformação linear e indicarmos com A^* a transformação induzida $\mathcal{A}_r(F) \to \mathcal{A}_r(E)$ (sem especificar na notação o grau r) então, para $\omega \in \mathcal{A}_r(F)$ e $\bar{\omega} \in \mathcal{A}_s(F)$ arbitrárias, teremos

$$A^*(\omega \wedge \bar{\omega}) = (A^*\omega) \wedge (A^*\bar{\omega}).$$

Novamente, esta relação é evidente quando ω e $\bar{\omega}$ são decomponíveis e o caso geral se reduz imediatamente a este.

Seja E um espaço vetorial de dimensão m. A soma direta

$$\Lambda(E^*) = \mathbb{R} \oplus E^* \oplus \mathcal{A}_2(E) \oplus \cdots \oplus \mathcal{A}_m(E)$$

é um espaço vetorial de dimensão 2^m (soma dos coeficientes binomiais $\binom{m}{r}$ quando r varia de 0 a m), ao qual o produto $\omega \wedge \bar{\omega}$, acima definido, pode ser obviamente estendido, de modo a dotar $\Lambda(E^*)$ de uma multiplicação associativa (não-comutativa). Com essa operação, $\Lambda(E^*)$ torna-se uma álgebra, conhecida como a *Álgebra de Grassmann* do espaço vetorial E^*.

2 Formas diferenciais

As formas diferenciais de grau r, que estudaremos agora, serão os integrandos das integrais de superfície. Elas generalizam as formas diferenciais de grau 1, consideradas no Capítulo IV, não apenas por serem as suas análogas de dimensão superior, como por estarem definidas em superfícies, em vez de somente abertos do espaço euclidiano.

Encerrado o prelúdio algébrico feito no parágrafo 1, voltaremos a usar a notação $\{dx_1, \ldots, dx_m\}$ para indicar a base canônica de $(\mathbb{R}^m)^*$, dual da base $\{e_1, \ldots, e_m\} \subset \mathbb{R}^m$, onde $e_1 = (1, 0, \ldots, 0)$, etc.

Para cada conjunto $I = \{i_1 < \cdots < i_r\} \subset \{1, 2, \ldots, m\}$, escreveremos

$$dx_1 = dx_{i_1} \wedge \cdots \wedge dx_{i_r}\,.$$

As formas r-lineares alternadas dx_I constituem a base canônica do espaço vetorial $\mathcal{A}_r(\mathbb{R}^m)$. Dada uma lista de r vetores $v_1, \ldots, v_r \in \mathbb{R}^m$, obtemos uma matriz $a = (a_{ij})$, com m linhas e r colunas, na qual a j-ésima coluna é o vetor $v_j = (a_{1j}, \ldots, a_{mj}$. Neste caso,

$$dx_I(v_1, \ldots, v_r) = \det(a_I),$$

onde a_I é a matriz $r \times r$ obtida de a selecionando-se as linhas cujos índices pertencem ao conjunto I. Geometricamente, $dx_I(v_1, \ldots, v_r)$ é o volume r-dimensional (orientado) da projeção do paralelepípedo $[v_1, \ldots, v_r]$ sobre o subespaço r-dimensional de $\mathbb{R}^m$ que tem $\{e_{i_1}, \ldots, e_{i_r}\}$ como base positiva. (Veja Exercício 2.2.)

Consideremos $m = 3$. A base canônica de $(\mathbb{R}^3)^* = \mathcal{A}_1(\mathbb{R}^3)$ é formada pelos funcionais dx, dy, dz, enquanto que $\{dx \wedge dy, dy \wedge dz, dx \wedge dz\}$ é a base canônica de $\mathcal{A}_2(\mathbb{R}^3)$. Dados os vetores $u = (a_1, a_2, a_3)$ e $v = (b_1, b_2, b_3)$ em $\mathbb{R}^3$, temos $(dx \wedge dy)(u, v) = a_1b_2 - a_2b_1 =$ área orientada da projeção do paralelogramo $[u, v]$ sobre o plano x, y. Também $(dy \wedge dz)(u, v) = a_2b_3 - a_3b_2$ e $(dx \wedge dz)(u, v) = a_1b_3 - a_3b_1$ têm interpretações

análogas. Note que $\{dx \wedge dy, dy \wedge dz, dz \wedge dx\}$, por exemplo, também é uma base de $\mathcal{A}_2(\mathbb{R}^3)$ mas estamos chamando canônica a base anterior, pois em $dx_I = dx_{i_1} \wedge \cdots \wedge dx_{i_r}$ os índices $i_1 < \cdots < i_r$ são sempre tomados na ordem crescente. Finalmente, a base canônica de $\mathcal{A}_3(\mathbb{R}^3)$ é formada pelo único elemento $dx \wedge dy \wedge dz$. Dados $u, v, w \in \mathbb{R}^3$, $(dx \wedge dy \wedge dz)(u, v, w)$ é o valor do determinante da matriz 3×3 cujas colunas são os vetores u, v, w.

Uma *forma diferencial de grau* r num aberto $U \subset \mathbb{R}^m$ é uma aplicação $\omega\colon U \to \mathcal{A}_r(\mathbb{R}^m)$. A cada ponto $x \in U$, ω faz corresponder a forma r-linear alternada $\omega(x) = \sum_I a_I(x) dx_I$. Assim, a forma diferencial ω determina (e é determinada por) funções $a_I\colon U \to \mathbb{R}$, chamadas as *coordenadas* de ω. Para cada subconjunto $I = \{i_1 < \cdots < i_r\} \subset \{1, 2, \ldots, m\}$, e cada ponto $x \in U$, temos $a_I(x) = \omega(x) \cdot (e_{i_1}, \ldots, e_{i_r})$.

Exemplo 11. Seja $U \subset \mathbb{R}^3$ um aberto. As formas diferenciais em U são:

1º) As funções $f\colon U \to \mathbb{R}$ (formas de grau zero).

2º) As formas de grau 1, $\omega = adx + bdy + cdz$, onde a, b, c são funções reais definidas em U. Toda forma de grau 1 em U pode ser identificada a um campo de vetores em U, com as mesmas coordenadas.

3º) As formas de grau 2, $\omega = adx \wedge dy + bdy \wedge dz + cdx \wedge dz$. Em geral, uma forma de grau 2 num aberto $U \subset \mathbb{R}^m$ é definida por $m(m-1)/2$ coordenadas, de modo que não corresponde necessariamente a um campo de vetores. No caso particular de $\mathbb{R}^3$, temos $3(3-1)/2 = 3$, por isso é possível associar a ω o campo de vetores de coordenadas a, b, c. Isto faz com que na Análise em $\mathbb{R}^3$ não haja necessidade de Álgebra Multilinear.

4º) As formas de grau 3, $\omega = adx \wedge dy \wedge dz$, que também correspondem a funções $a\colon U \to \mathbb{R}$.

Diz-se que a forma ω é *de classe* C^k quando cada coordenada a_I é uma função de classe C^k. A fim de que $\omega \in C^k$, é necessário e suficiente que, dados arbitrariamente $v_1, \ldots, v_r \in \mathbb{R}^m$, a função $x \mapsto \omega(x)(v_1, \ldots, v_r)$ seja de classe C^k no aberto U. Com efeito, pondo $c_I = dx_I(v_1, \ldots, v_r)$, temos $\omega(x) \cdot (v_1, \ldots, v_r) = \sum_I a_I(x) \cdot c_I$, logo a condição é necessária. Reciprocamente, como $a_I(x) = \omega(x) \cdot (e_{i_1}, \ldots, e_{i_r})$, a condição é suficiente.

Uma *forma diferencial de grau* r *numa superfície* m-dimensional

$M \subset \mathbb{R}^n$ é uma aplicação

$$\omega\colon x \in M \mapsto \omega(x) \in \mathcal{A}_r(T_xM),$$

que associa a cada ponto $x \in M$ uma forma r-linear alternada $\omega(x)$ no espaço vetorial tangente T_xM.

Se $r = 0$, uma forma diferencial de grau zero em M é simplesmente uma função real $\omega\colon M \to \mathbb{R}$.

Seja $\varphi\colon U_0 \to U$ uma parametrização de um aberto $U \subset M$. Em cada ponto $x = \varphi(u) \in U$, temos a base $\left\{\dfrac{\partial\varphi}{\partial u_1}(u), \ldots, \dfrac{\partial\varphi}{\partial u_m}(u)\right\} \subset T_xM$. Usaremos a notação $\{du_1, \ldots, du_m\} \subset (T_xM)^*$ para indicar a base dual.

Na realidade, $du_1, \ldots, du_m$ são formas diferenciais de grau 1 em U: para cada ponto $x \in U$, temos os funcionais lineares $du_1(x), \ldots, du_m(x)$, pertencentes a $(T_xM)^*$. Usamos a notação simplificada du_i, em vez de $du_i(x)$, pois não há perigo de confusão.

Em cada ponto $x = \varphi(u) \in U$, as r-formas $du_I = du_{i_1} \wedge \cdots \wedge du_{i_r}$, com $I = \{i_1 < \cdots < i_r\}$, constituem uma base de $\mathcal{A}_r(T_xM)$. Dada uma forma diferencial ω, de grau r em M, podemos então escrever, para cada ponto $x = \varphi(u) \in U$:

$$\omega(x) = \omega(\varphi(u)) = \sum_I a_I(u)du_I\,.$$

Assim, a forma ω define, para cada parametrização $\varphi\colon U_0 \to U$ em M, as funções $a_I\colon U_0 \to \mathbb{R}$, em número de $\binom{m}{r}$. Elas são chamadas as *coordenadas* da forma ω relativamente à parametrização φ. Pondo, por simplicidade, $w_i = \dfrac{\partial\varphi}{\partial u_i}(u)$, temos $a_I(u) = \omega(\varphi(u)) \cdot (w_{i_1}, \ldots, w_{i_r})$, para todo $u \in U_0$.

Seja agora $\psi\colon V_0 \to V$ outra parametrização em M, com $U \cap V \neq \varnothing$. Para todo $x = \varphi(u) = \psi(v) \in U \cap V$, temos as bases duais

$$\left\{\frac{\partial\psi}{\partial v_1}(v), \ldots, \frac{\partial\psi}{\partial v_m}(v)\right\} \subset T_xM \quad \text{e} \quad \{dv_1, \ldots, dv_m\} \subset (T_xM)^*,$$

que se relacionam com as bases determinadas por φ do seguinte modo:

$$\frac{\partial\varphi}{\partial u_j} = \sum_{i=1}^m \frac{\partial v_i}{\partial u_j} \cdot \frac{\partial\psi}{\partial v_i} \quad \text{e} \quad dv_i = \sum_{j=1}^m \frac{\partial v_i}{\partial u_j}\, du_j\,.$$

Nas igualdades acima, $(\partial v_i/\partial u_j)$ é a matriz jacobiana da mudança de parametrização $\psi^{-1} \circ \varphi$, calculada no ponto u, a derivada $\partial\varphi/\partial u_j$ é tomada no ponto u e $\partial\psi/\partial v_i$ no ponto $v = \psi^{-1}(\varphi(u))$.

A parametrização ψ determina em $\mathcal{A}_r(T_xM)$ a base constituida pelas r-formas $dv_I = dv_{i_1} \wedge \cdots \wedge dv_{i_r}$.

Pelo que vimos no parágrafo 1, se $x = \varphi(u) = \psi(v) \in U \cap V$ e $\omega(x) = \sum\limits_J a_J(u)du_J = \sum\limits_I b_I(v)dv_I$ então

$$(*) \qquad a_J(u) = \sum_I \det\left(\frac{\partial v_I}{\partial u_J}\right) b_I(v),$$

onde $(\partial v_I/\partial u_J)$ é a matriz $r \times r$ formada pelos elementos $\dfrac{\partial v_i}{\partial u_j}(v)$ tais que $i \in I$ e $j \in J$.

Nos termos do Cálculo Tensorial clássico, uma forma diferencial de grau r numa superfície M seria uma aplicação que a cada parametrização $\varphi\colon U_0 \to U$ em M associaria $\binom{m}{r}$ funções $a_J\colon U_0 \to \mathbb{R}$, chamadas as *coordenadas* da forma relativamente a φ, de tal modo que se à parametrização $\psi\colon V_0 \to V$ correspondem as funções $b_I\colon V_0 \to \mathbb{R}$ e $\varphi(u) = \psi(v)$, então valem as igualdades (*), chamadas as "fórmulas de mudança de coordenadas".

Convém destacar o importante caso particular em que $r = m$, ou seja, o grau da forma é igual à dimensão da superfície. Neste caso, a forma tem apenas uma coordenada em cada parametrização. Assim, para todo ponto $x = \varphi(u) = \psi(v) \in U \cap V$, temos

$$\omega(x) = a(u)du_1 \wedge \cdots \wedge du_m = b(v)dv_1 \wedge \cdots \wedge dv_m\,,$$

onde as funções $a\colon U_0 \to \mathbb{R}$ e $b\colon V_0 \to \mathbb{R}$ cumprem

$$(**) \qquad a(u) = \det\left(\frac{\partial v_i}{\partial u_j}\right) \cdot b(v).$$

Acima, $u \in \varphi^{-1}(U \cap V)$, $v = (\psi^{-1} \circ \varphi)(u)$ e $\det.(\partial v_i/\partial u_j)$ é o determinante jacobiano do difeomorfismo $\psi^{-1} \circ \varphi$, calculado no ponto u.

Seja M uma superfície de classe C^s. Uma forma diferencial ω, de grau r em M, diz-se *de classe* C^k $(k < s)$ quando M pode ser coberta por imagens U de parametrizações $\varphi\colon U_0 \to U$, de classe C^k, relativamente às quais se tem $\omega = \Sigma\, a_I du_I$, as funções $a_I\colon U_0 \to \mathbb{R}$ sendo todas de

classe C^k. As fórmulas de mudança (*) mostram que se as coordenadas de ω numa parametrização $\psi \in C^s$ são funções de classe C^k $(k < s)$ então elas ainda serão de classe C^k em qualquer outra parametrização $\varphi \in C^s$ com $\mathrm{Im}(\varphi) \cap \mathrm{Im}(\psi) \neq \varnothing$. Uma forma de grau zero e classe C^k em M é simplesmente uma função $f\colon M \to \mathbb{R}$, de classe C^k.

As fórmulas de mudança de coordenadas exprimem a "invariância" das formas diferenciais. Seu caso mais simples (**), por exemplo, assegura que a integral de uma forma de grau m esteja bem definida sobre uma superfície orientada de dimensão m, como será visto agora.

Consideraremos apenas um caso particular dessa noção de integral de superfície. A situação geral será reduzida a esta, mediante as partições da unidade, no parágrafo 6.

O *suporte* de uma forma diferencial ω numa superfície M é o fecho (relativamente a M) do conjunto dos pontos $x \in M$ tais que $\varphi(x) \neq 0$. Escreveremos supp.ω para representar o suporte de ω. Assim, $x \in$ supp.ω significa que toda vizinhança de x contém pontos y tais que $\omega(y) \neq 0$. Contrariamente, $x \notin$ supp.ω quer dizer que existe uma vizinhança V de x em M tal que $\omega(y) = 0$ para todo $y \in V$. Por definição, o suporte de ω é sempre um subconjunto fechado de M. Se a forma ω é contínua e $\omega(x) \neq 0$ então ω é diferente de zero numa vizinhança de x, logo x pertence ao interior de supp. ω (relativamente a M). Noutras palavras, se $\omega \in C^0$ e x é um ponto da fronteira de supp. ω então $\omega(x) = 0$, embora seja $x \in$ supp. ω.

Definiremos agora a integral de uma forma contínua ω, de grau m, com suporte compacto, sobre uma superfície orientada M, de dimensão m, no caso particular em que supp. ω está contido na imagem U de uma parametrização positiva $\varphi\colon U_0 \to U$.

Em termos dessa parametrização, tem-se $\omega(x) = a(u)du_1 \wedge \cdots \wedge du_m$ para todo $x = \varphi(u) \in U$, onde a função contínua $a\colon U_0 \to \mathbb{R}$ tem suporte compacto, igual a $\varphi^{-1}(\text{supp.}\,\omega)$. Por definição, pomos

$$\int_M \omega = \int_K a(u)du,$$

onde $K \subset \mathbb{R}^m$ é qualquer compacto J-mensurável, contido em U_0 e contendo o suporte de a. (Por exemplo: cubra o compacto supp. a com um número finito de cubos fechados, contidos em U_0, e chame de K a reunião desses cubos.) Se quisermos, poderemos retirar a exigência $K \subset U_0$ (mantendo supp. $a \subset K$!), desde que consideremos a como uma função contínua definida em K, igual a zero nos pontos de $K - U_0$.

Evidentemente, a escolha do conjunto K nas condições acima não afeta o valor da integral $\int_K a(u)du$.

Resta entretanto mostrar que $\int_M \omega$, definida da maneira acima, não depende da escolha da parametrização positiva φ.

Com efeito, seja $\psi\colon V_0 \to V$ outra parametrização positiva em M, com supp. $\omega \subset V$ e $\omega(x) = b(v)dv_1 \wedge \cdots \wedge dv_m$ para todo $x = \psi(v) \in v$. A função $b\colon V_0 \to \mathbb{R}$ é contínua, seu suporte é igual a $\psi^{-1}(\text{supp.}\,\omega)$ e, para todo $u \in \varphi^{-1}(U \cap V)$, tem-se $a(u) = J(u) \cdot b(v)$, onde $J(u)$ é o valor do determinante jacobiano de $\psi^{-1} \circ \varphi$ no ponto u e $v = \psi^{-1}\varphi(u)$. Observemos que $J(u) > 0$ para todo $u \in \varphi^{-1}(U \cap V)$. Para calcular $\int_M \omega$ em termos da parametrização φ, usemos um compacto J-mensurável K, com $\varphi^{-1}(\text{supp.}\,\omega) \subset K \subset \varphi^{-1}(U \cap V)$ e, para calcular $\int_M \omega$ em termos de ψ, usemos o compacto J-mensurável $L = \psi^{-1}\varphi(K)$. Pelo Teorema de Mudança de Variáveis,

$$\int_L b(v)dv = \int_K b(\psi^{-1}\varphi(u)) \cdot J(u)du = \int_K a(u)du.$$

Isto mostra que a integral de superfície $\int_M \omega$ está bem definida, sempre que ω seja uma forma diferencial contínua de grau m, com suporte compacto, contido numa vizinhança parametrizada em uma superfície orientada m-dimensional M.

No parágrafo 6 definiremos a integral $\int_M \omega$ sob condições mais gerais para a forma ω.

Mostraremos agora que toda aplicação diferenciável $f\colon M \to N$ induz uma transformação linear $\omega \mapsto f^*\omega$, a qual leva formas de grau r na superfície N em formas de grau r na superfície M. Essa transformação é uma das coisas que tornam úteis as formas diferenciais como instrumentos para o estudo das aplicações de uma superfície noutra.

Seja, então, $f\colon M \to N$ de classe C^k $(k \geq 1)$. Dada a forma ω, de grau r sobre N, definiremos a forma $f^*\omega$, de grau r sobre M, pondo para cada $x \in M$ e cada r-lista de vetores $w_1, \ldots, w_r \in T_xM$,

$$[(f^*\omega)(x)](w_1, \ldots, w_r) = \omega(f(x)) \cdot (f'(x) \cdot w_1, \ldots, f'(x) \cdot w_r).$$

Como sabemos, a aplicação $f\colon M \to N$ possui, em cada ponto $x \in M$, uma derivada, que é uma transformação linear $f'(x)\colon T_xM \to T_{f(x)}N$. Esta, por sua vez, induz uma transformação linear $[f'(x)]^*\colon \mathcal{A}_r(T_{f(x)}N) \to \mathcal{A}_r(T_xM)$, como foi visto no parágrafo 1. De acordo com a definição dada ali, temos $(f^*\omega)(x) = [f'(x)]^* \cdot \omega(f(x))$.

No caso de uma forma de grau zero, isto é, de uma função $g\colon N \to \mathbb{R}$, temos $f^*(g) = g \circ f$.

As propriedades abaixo resultam imediatamente dos fatos análogos sobre formas r-lineares alternadas:

1. $f^*(\omega + \bar{\omega}) = f^*\omega + f^*\bar{\omega}$;
2. $f^*(c \cdot \omega) = c \cdot f^*\omega, \quad c \in \mathbb{R}$;
3. $f^*(\omega \wedge \bar{\omega}) = f^*\omega \wedge f^*\bar{\omega}$;
4. $(g \circ f)^*\omega = f^*(g^*\omega)$.

No caso 4), temos três superfícies M, N, P, duas aplicações $f\colon M \to N$, $g\colon N \to P$, e uma forma ω em P. Como caso particular de 3), tem-se o produto $\lambda{\cdot}\omega$ de uma forma ω em N por uma função $\lambda\colon N \to \mathbb{R}$. Então $f^*(\lambda{\cdot}\omega) = f^*\lambda{\cdot} f^*\omega = (\lambda{\circ} f){\cdot} f^*\omega$. Resulta de 4) que se $f\colon M \to N$ é um difeomorfismo, cujo inverso é $g\colon N \to M$, então $\omega \mapsto f^*\omega$ é uma bijeção das formas em N sobre as formas em M, cuja inversa é a aplicação $\alpha \mapsto g^*\alpha$.

Em termos de coordenadas:

Sejam $\varphi\colon U_0 \to U \subset M$, $\psi\colon V_0 \to V \subset N$ parametrizações de classe C^k, com $f(U) \subset V$, e $\omega(y) = \sum_I b_I(v)dv_I$ para todo $y = \psi(v) \in V$. Então, para qualquer $x = \varphi(u) \in U$, temos $(f^*\omega)(x) = \sum_J a_J(u)du_J$, onde as coordenadas $a_J\colon U_0 \to \mathbb{R}$ são determinadas a partir das coordenadas $b_I\colon V_0 \to \mathbb{R}$ pelas igualdades

$$(*) \qquad a_J(u) = \sum_I \det\left(\frac{\partial v_I}{\partial u_J}(u)\right) b_I(v),$$

onde $u \in U_0$, $v = (\psi^{-1} \circ f \circ \varphi)(u)$, $(\partial v_i/\partial u_j)$ é a matriz jacobiana da derivada $\psi^{-1} \circ f \circ \varphi\colon U_0 \to V_0$ no ponto u, enquanto que $\det(\partial v_I/\partial u_J)$ é o determinante da submatriz $r \times r$ obtida tomando-se os $\partial v_i/\partial u_j$ com $i \in I$ e $j \in J$.

A fórmula acima resulta do cálculo análogo que fizemos no parágrafo 1 para o caso algébrico, se observarmos que a matriz da transformação linear $f'(x)\colon T_xM \to T_yN$, $y = f(x)$, relativamente às bases

$$\left\{\frac{\partial \varphi}{\partial u_1}(u), \cdots, \frac{\partial \varphi}{\partial u_m}(u)\right\} \subset T_xM \text{ e } \left\{\frac{\partial \psi}{\partial v_1}(v), \cdots, \frac{\partial \psi}{\partial v_n}(v)\right\} \subset T_yN,$$

onde $y = \psi(v)$, é precisamente a matriz $(\partial v_i/\partial u_j)$.

Em particular, para cada I, temos

$$f^*(dv_I) = \sum_J \det\left(\frac{\partial v_I}{\partial u_J}\right) du_j\,.$$

Assim, para passar de $\omega = \Sigma\, b_I dv_I$ para $f^*\omega = \Sigma\, a_j du_J$, basta substituir cada dv_i por $\sum_J \det\left(\frac{\partial v_I}{\partial u_J}\right) du_J$. Isto significa que, do ponto de vista computacional, a operação $\omega \mapsto f^*\omega$ é simplesmente uma mudança de variáveis $y = f(x)$, ou $v = (\psi^{-1} f\varphi)(u)$.

Segue-se da fórmula de passagem das coordenadas b_I para as a_J que, se ω é uma forma de grau ≥ 1 e classe C^k em N e $f\colon M \to N$ é uma aplicação de classe C^{k+1}, então $f^*\omega$ é uma forma de classe C^k em M.

Observações: 1. Nas fórmulas (*), a fim de obter a expressão de cada a_J como uma função em U_0, podemos substituir v por $(\psi^{-1} f\varphi)(u)$, resultando assim as funções $\bar{b}_I(u) = b_I((\psi^{-1} f\varphi)(u)) = b_I(v)$ no segundo membro.

2. A forma $f^*\omega$ chama-se às vezes o *pull-back* da forma ω por meio de f.

Exemplo 12. Como primeiro exemplo de forma induzida por uma aplicação, seja ω uma forma de grau r na superfície M. Seja $\varphi\colon U_0 \to U$ uma parametrização em M. Em cada ponto $x = \varphi(u) \in U$, os vetores $\frac{\partial\varphi}{\partial u_1}(u), \ldots, \frac{\partial\varphi}{\partial u_m}(u)$ constituem uma base do espaço vetorial tangente T_xM, cuja dual é $\{du_1, \ldots, du_m\} \subset (T_xM)^*$. Escrevendo $du_I = du_{i_1} \wedge \cdots \wedge du_{i_r}$, com $I = \{i_1 < \cdots < i_r\}$, temos $\omega(x) = \Sigma\, a_I(u) du_I$. Se considerarmos φ como uma aplicação diferenciável $\varphi\colon U_0 \to M$, afirmamos que, para todo $u = \varphi^{-1}(x) \in U_0$, vale $(\varphi^*\omega)(u) = \sum_I a_I(u) dx_I$. Aqui $dx_I = dx_{i_1} \wedge \cdots \wedge dx_{i_r}$, onde $\{dx_1, \ldots, dx_m\}$ é a base canônica de $(\mathbb{R}^m)^*$, dual de $\{e_1, \ldots, e_m\} \subset \mathbb{R}^m$. Sabemos que $\varphi'(u) \cdot e_i = \frac{\partial\varphi}{\partial u_i}(x)$. Portanto:

$$(\varphi^* du_i)(e_j) = du_i(\varphi'(u) \cdot e_j) = du_i\left(\frac{\partial\varphi}{\partial u_j}\right) = \begin{cases} 0 \text{ se } i \neq j \\ 1 \text{ se } i = j \end{cases}.$$

Isto mostra que $\varphi^* du_i = dx_i$. Mais geralmente, se $I = \{i_1 < \cdots < i_r\}$, vale:

$$[(\varphi^*\omega)(u)](e_{i_1}, \ldots, e_{i_r}) = \omega(x)\left(\frac{\partial\varphi}{\partial u_{i_1}}(u), \ldots, \frac{\partial\varphi}{\partial u_{i_r}}(u)\right) = a_I(u).$$

Logo $\varphi^*\omega = \Sigma\, a_I dx_I$, como queríamos mostrar.

Exemplo 13. Outro exemplo de forma induzida (com as mesmas notações do Ex. 12) é o seguinte. Se $\omega|U = \Sigma\, a_I du_I$ na parametrização φ, $\omega|V = \Sigma\, b_J dv_J$ na parametrização ψ e $\xi = \psi^{-1} \circ \varphi \colon \varphi^{-1}(U \cap V) \to \psi^{-1}(U \cap V)$ é a mudança de parametrização, então $\xi^*(\Sigma\, b_J dx_J) = \Sigma\, a_I dx_I$. Com efeito, temos $\varphi = \psi \circ \xi$, $\varphi^*\omega = \Sigma\, a_I dx_I$ e $\psi^*\omega = \Sigma\, b_J dx_J$. Logo,

$$\Sigma\, a_I dx_I = \varphi^*\omega = \xi^*\psi^*\omega = \xi^*(\Sigma\, b_J dx_J).$$

Exemplo 14. Sejam M, N superfícies com $M \subset N$. Indiquemos com $i \colon M \to N$ a aplicação de inclusão, $i(x) = x$ para todo $x \in M$. Se ω é uma forma em N, $i^*\omega$ é uma forma em M, chamada a *restrição* de ω a M. Para todo $x \in M$, temos $T_xM \subset T_xN$ e $(i^*\omega)(x)$ é a restrição de $\omega(x)$ a r-listas de vetores pertencentes a T_xM. Por exemplo, se $N = \mathbb{R}^3$, $M = \{(0, y, z) \in \mathbb{R}^3\}$ e $\omega = dx \wedge dy$, então $i^*\omega = 0$. Outro caso particular de restrição: seja $\omega = ady \wedge dz + bdz \wedge dx + cdx \wedge dy$ uma forma contínua de grau 2, definida no aberto $A \subset \mathbb{R}^3$. Seja M uma superfície orientada (bidimensional) contida em A. Em termos de uma parametrização $\varphi \colon U_0 \to U$ em M, com $\varphi(u, v) = (x(u, v), y(u, v), z(u, v))$, a restrição $i^*\omega$ assume a expressão

$$i^*\omega = \left[\bar{a} \cdot \det. \frac{\partial(y, z)}{\partial(uv)} + \bar{b} \cdot \det. \frac{\partial(z, x)}{\partial(u, v)} + \bar{c} \cdot \det. \frac{\partial(x, y)}{\partial(u, v)}\right] du \wedge dv,$$

onde $\bar{a} = a \circ \varphi$, $\quad \dfrac{\partial(y, z)}{\partial(u, v)} = \begin{pmatrix} \dfrac{\partial y}{\partial u} & \dfrac{\partial z}{\partial u} \\ \\ \dfrac{\partial y}{\partial v} & \dfrac{\partial z}{\partial v} \end{pmatrix}$, etc.

O coeficiente de $du \wedge dv$ é o produto interno do vetor $F = (\bar{a}, \bar{b}, \bar{c})$ pelo vetor normal $N(u, v) = \dfrac{\partial \varphi}{\partial u} \times \dfrac{\partial \varphi}{\partial v}$ (vide Exemplo 4). Por definição, põe-se $\int_M \omega = \int_M i^*\omega$ se $i^*\omega$ tem suporte compacto, contido em U). Então

$$\int_M \omega = \int_{U_0} \langle F, N \rangle du dv,$$

como nos livros de Cálculo.

De um modo geral, a definição que demos para $\int_M \omega$, onde ω tem suporte compacto, contido na imagem da parametrização positiva

$\varphi\colon U_0 \to U$, permite que escrevamos (considerando U e U_0 como superfícies m-dimensionais):

$$\int_M \omega = \int_U \omega = \int_{U_0} \varphi^* \omega.$$

Exemplo 15. (A forma elemento de volume.) Seja M uma superfície orientada. Para todo ponto $x \in M$, o espaço vetorial tangente T_xM possui um produto interno (já que é um subespaço de $\mathbb{R}^n$) e é orientado: se $\varphi\colon U_0 \to U$ é uma parametrização positiva em M, em cada ponto $x = \varphi(u) \in U$ a base $\left\{\dfrac{\partial\varphi}{\partial u_1}(u), \ldots, \dfrac{\partial\varphi}{\partial u_m}(u)\right\}$ define a orientação positiva de T_xM. Logo podemos introduzir uma forma diferencial ω, de grau $m = \dim M$ pondo, para cada $x \in M$, $\omega(x) =$ elemento de volume do espaço vetorial T_xM, nos termos do Exemplo 9. Para $x \in M$ e $w_1, \ldots, w_m \in T_xM$, temos $\omega(x)(w_1, \ldots, w_m) =$ volume orientado do paralelepípedo $[w_1, \ldots, w_m]$. Se $M \in C^k$ e $\varphi\colon U_0 \to U$ é uma parametrização de classe C^k em M então, para cada $x = \varphi(u) \in U$, temos $\omega(x) = \sqrt{g(u)}du_1 \wedge \cdots \wedge du_m$, onde $g(u) = \det(g_{ij}(u))$, $g_{ij}(u) = \langle\dfrac{\partial\varphi}{\partial u_i}(u), \dfrac{\partial\varphi}{\partial u_j}(u)\rangle$, e $\{du_1, \ldots, du_m\} \subset (T_xM)^*$ é a base dual de $\left(\dfrac{\partial\varphi}{\partial u_i}(u)\right)$. Isto mostra que a forma ω é de classe C^{k-1}. No caso particular em que $m = 2$, escrevemos $(u, v) \in U_0$, em vez de (u_1, u_2). Nesta situação, é tradicional pôr $E = \langle\dfrac{\partial\varphi}{\partial u}, \dfrac{\partial\varphi}{\partial u}\rangle$, $F = \langle\dfrac{\partial\varphi}{\partial u}, \dfrac{\partial\varphi}{\partial v}\rangle$ e $G = \langle\dfrac{\partial\varphi}{\partial v}, \dfrac{\partial\varphi}{\partial v}\rangle$. Então o *elemento de área* da superfície M se escreve como $\omega = \sqrt{EG - F^2}du \wedge dv$. Quando $m = 1$, uma parametrização positiva da curva orientada M tem a forma $\varphi\colon J_0 \to J$, onde J_0 é um intervalo aberto da reta. Em cada ponto $x = \varphi(t) \in J$, temos $\omega(x) = |\varphi'(t)|dt$, onde $\varphi'(t) \in T_xM$ é o vetor básico associado à parametrização φ e $\{dt\} \subset (T_xM)^*$ é a base dual de $\{\varphi'(t)\}$. A forma $\omega = |\varphi'(t)|dt$ chama-se o *elemento de arco* da curva M. Para todo vetor tangente $w \in T_xM$, podemos escrever $w = \alpha \cdot \varphi'(t)$, donde $|\alpha| = |w|/|\varphi'(t)|$. Então $\omega(x) \cdot w = |\varphi'(t)|dt(w) = |\varphi'(t)| \cdot \alpha = \pm|w|$, o sinal sendo positivo se $\alpha > 0$, isto é, se w tem a mesma orientação de $\varphi'(t)$, e negativo no caso contrário. Mencionemos, finalmente, que se $U \subset \mathbb{R}^m$ é um aberto, munido de sua orientação natural, o elemento de volume de U é a forma diferencial $\omega = dx_1 \wedge \cdots \wedge dx_m$.

Exemplo 16. (Elemento de volume de uma hiperfície.) Seja $M \subset \mathbb{R}^{m+1}$ uma hiperfície orientada, de classe C^k. Existe um campo unitário de vetores normais $v\colon M \to \mathbb{R}^{m+1}$, de classe C^{k-1}, tal que $\{w_1, \ldots, w_m\} \subset T_xM$ é uma base positiva se, e somente se, $\det(v(x), w_1, \ldots, w_m) > 0$. [Aqui, $(v(x), w_1, \ldots, w_m)$ é a matriz $(m+1) \times (m+1)$ cujas colunas são os vetores indicados.] Se $w_1, \ldots, w_m$ são m vetores quaisquer do espaço vetorial tangente T_xM, como $v(x)$ é perpendicular a T_xM e tem comprimento 1, o volume do paralelepípedo $[w_1, \ldots, w_m]$ é igual a $|\det(v(x), w_1, \ldots, w_m)|$. Se ω é o elemento de volume de M, temos então, para todo $x \in M$ e quaisquer $w_1, \ldots, w_m \in T_xM$:

$$\omega(x)(w_1, \ldots, w_m) = \det(v(x), w_1, \ldots, w_m),$$

já que o sinal do determinante coincide com aquele da definição de ω. Desenvolvendo esse determinante ao longo de sua primeira coluna $v(x) = (v_1(x), \ldots, v_{m+1}(x))$, obtemos

$$\omega(x)(w_1, \ldots, w_m) = \sum_{i=1}^{m+1} (-1)^{i+1}\, v_i(x) \cdot A_i\,,$$

onde A_i é o determinante da matriz $m \times m$, obtida omitindo-se a i-ésima linha da matriz $(m+1) \times m$ cujas colunas são $w_1, \ldots, w_m$. Se escrevermos $dx_1 \wedge \cdots \wedge \widehat{dx_i} \wedge \cdots \wedge dx_{m+1}$ em vez de $dx_1 \wedge \cdots \wedge dx_{i-1} \wedge dx_{i+1} \wedge \cdots \wedge dx_{m+1}$, veremos imediatamente que

$$A_i = (dx_1 \wedge \cdots \wedge \widehat{dx_i} \wedge \cdots \wedge dx_{m+1})(w_1, \ldots, w_m).$$

Por conseguinte

$$\omega(x) \cdot (w_1, \ldots, w_m) =$$

$$= \left[\sum_{i=1}^{m+1} (-1)^{i+1}\, v_i(x) dx_1 \wedge \cdots \wedge \widehat{dx_i} \wedge \cdots \wedge dx_{m+1}\right] (w_1, \ldots, w_m).$$

Isto significa que

$$\omega(x) = \Sigma(-1)^{i+1}\, v_i(x) \cdot dx_1 \wedge \cdots \wedge \widehat{dx_i} \wedge \cdots \wedge dx_{m+1}\,.$$

Para calcular a normal unitária $v(x)$ em casos particulares, é conveniente observar que se $\varphi\colon U_0 \to U$ é uma parametrização positiva em M e $x = \varphi(u)$, então $v(x) = N(x)/|N(x)|$, onde $N(x) = \dfrac{\partial \varphi}{\partial u_1}(u) \times \cdots \times$

$\dfrac{\partial\varphi}{\partial u_m}(u)$, pois este produto vetorial é perpendicular ao espaço tangente T_xM e tem o sentido de $v(x)$. Em termos dos parâmetros $u_1, \ldots, u_m$, temos, na realidade,

$$\omega(x) = \left|\frac{\partial\varphi}{\partial u_1}(u) \times \cdots \times \frac{\partial\varphi}{\partial u_m}(u)\right| du_1 \wedge \cdots \wedge du_m\,,\ x = \varphi(u),$$

pois o comprimento do produto vetorial de m vetores é o volume do paralelepípedo que tem esses vetores como arestas e sabemos que $\omega(x) = a(u)du_1 \wedge \cdots \wedge du_m$, onde

$$a(u) = \omega(x) \cdot \left(\frac{\partial\varphi}{\partial u_1}(u), \ldots, \frac{\partial\varphi}{\partial u_m}(u)\right) =$$

$$= \text{volume do paralelepípedo}\left[\frac{\partial\varphi}{\partial u_1}(u), \ldots, \frac{\partial\varphi}{\partial u_m}(u)\right].$$

Em particular, seja $M = \{x \in \mathbb{R}^{m+1}; |x-a| = r\}$ a esfera de centro $a \in \mathbb{R}^{n+1}$ e raio $r > 0$. Então, para cada ponto $x \in M$, podemos tomar $v(x) = \dfrac{x-a}{|x-a|} = \dfrac{x-a}{r}$ (normal unitária exterior). Neste caso, obtemos

$$\omega = \sum_{i=1}^{m+1} (-1)^{i+1}\,\frac{x_i - a_i}{r}\, dx_1 \wedge \cdots \wedge \widehat{dx_i} \wedge \cdots \wedge dx_{m+1}\,.$$

A expressão acima define uma forma diferencial de grau m, cujo domínio é todo o espaço euclidiano $\mathbb{R}^{m+1}$. Sua restrição à esfera de centro $a = (a_1, \ldots, a_{m+1})$ e raio r é o elemento de volume dessa esfera. Quando $m = 3$, e S é a esfera de centro no ponto (a, b, c) e raio r em $\mathbb{R}^3$, o elemento de volume de S é a restrição a S da forma diferencial

$$\omega = \frac{x-a}{r}\, dy \wedge dz + \frac{y-b}{r}\, dz \wedge dx + \frac{z-c}{r}\, dx \wedge dy.$$

Suponhamos agora que $M = f^{-1}(c)$ seja a imagem inversa de um valor regular $c \in \mathbb{R}$ por uma função $f\colon U \to \mathbb{R}$, de classe C^k no aberto $U \subset \mathbb{R}^{m+1}$. Neste caso, a orientação de M pode ser definida pelo campo unitário de vetores normais $v(x) = \operatorname{grad} f(x)/|\operatorname{grad} f(x)|$, de modo que $v_i(x) = \dfrac{\partial f}{\partial x_i}\Big/\sqrt{\sum_i \left(\dfrac{\partial f}{\partial x_i}\right)^2}$, as derivadas parciais sendo calculadas no

ponto x. Então o elemento de volume de M é a forma

$$\omega = \sum_{i=1}^{m+1} (-1)^{i+1} \frac{\partial f/\partial x_i}{\sqrt{\Sigma(\partial f/\partial x_i)^2}} dx_1 \wedge \cdots \wedge \widehat{dx_i} \wedge \cdots \wedge dx_{m+1}\,,$$

restrita a M.

Finalmente, a superfície M pode ser o gráfico de uma função $g\colon U \to \mathbb{R}$, definida no aberto $U \subset \mathbb{R}^m$. Então $M = f^{-1}(0)$, onde $f\colon U \times \mathbb{R} \to \mathbb{R}$, definida $f(x,t) = t - g(x)$ tem 0 como valor regular. Neste caso, para $m = 2$, por exemplo, vemos que o elemento de área da superfície M, gráfico da função $z = g(x,y)$ é dado pela forma de grau 2

$$\omega = \frac{\dfrac{\partial g}{\partial x} dz \wedge dy + \dfrac{\partial g}{\partial y} dx \wedge dz + dx \wedge dy}{\sqrt{\left(\dfrac{\partial g}{\partial x}\right)^2 + \left(\dfrac{\partial g}{\partial y}\right)^2 + 1}}$$

pois $\operatorname{grad} f = \left(-\dfrac{\partial g}{\partial x}, -\dfrac{\partial g}{\partial y}, 1\right)$.

Uma forma ω de grau m numa superfície orientada m-dimensional M diz-se *positiva* quando, para todo ponto $x \in M$ e toda base positiva $\{w_1, \ldots, w_m\} \subset T_xM$, tem-se $\omega(x)(v_1, \ldots, v_m) > 0$.

Neste caso, dada qualquer parametrização positiva $\varphi\colon U_0 \to U$ em M, tem-se $\omega(x) = a(u) du_1 \wedge \cdots \wedge du_m$, com $a(u) > 0$ para todo $u = \varphi^{-1}(x) \in U_0$, pois $a(u) = \omega(x) \cdot \left(\dfrac{\partial \varphi}{\partial u_1}(u), \ldots, \dfrac{\partial \varphi}{\partial u_m}(u)\right)$ e a base $\left(\dfrac{\partial \varphi}{\partial u_i}(u)\right)$ é positiva. Reciprocamene, se $a(u) > 0$ para toda parametrização positiva φ, então, dada arbitrariamente uma base positiva de vetores $w_j = \sum_i \alpha_{ij} \dfrac{\partial \varphi}{\partial u_i} \in T_xM$, tem-se

$$\begin{aligned}\omega(w)(w_1, \ldots, w_m) &= \det(\alpha_{ij}) \cdot \omega(x) \left(\frac{\partial \varphi}{\partial u_1}, \ldots, \frac{\partial \varphi}{\partial u_m}\right) = \\ &= \det(\alpha_{ij}) \cdot a(u) > 0,\end{aligned}$$

logo a forma ω é positiva.

Em toda superfície orientada de classe C^k existe uma forma positiva de classe C^{k-1}, a saber, o elemento de volume.

Reciprocamente, se existe numa superfície m-dimensional M uma forma contínua ω, de grau m, tal que $\omega(x) \neq 0$ para todo $x \in M$ então M é orientável (e há uma orientação de M que torna ω positiva). Com efeito, seja $\mathcal{A}$ o conjunto das parametrizações $\varphi \colon U_0 \to U$ em M tais que, para todo $x = \varphi(u) \in U$ tem-se $\omega(x) = a(u)du_1 \wedge \cdots \wedge du_m$, com $a(u) > 0$. Pela continuidade de ω (e portanto de a), se U_0 é conexo e $a(u) > 0$ em algum ponto $u \in U_0$ então $a > 0$ em todos os pontos de U_0. Daí resulta imediatamente que $\mathcal{A}$ é um atlas. Além disso, se $\varphi \colon U_0 \to U$ e $\psi \colon V_0 \to V$ pertencem a $\mathcal{A}$, com $U \cap V \neq \varnothing$ então $a(u) = \det(\psi^{-1} \circ \varphi)(u) \cdot b(v)$, para todo $u \in \varphi^{-1}(U \cap V)$ e $v = (\psi^{-1} \circ \varphi)(u)$. Como $a(u) > 0$ e $b(v) > 0$, segue-se que $\det(\psi^{-1} \circ \varphi)(u) > 0$ para todo $u \in \varphi^{-1}(U \cap V)$, logo o atlas $\mathcal{A}$ é coerente.

Exemplo 17 (O elemento de ângulo sólido.) Definimos agora uma forma diferencial de grau m em $\mathbb{R}^{m+1} - \{0\}$, a qual desempenha um papel análogo ao da forma elemento de ângulo $\delta\theta$ em $\mathbb{R}^2 - \{0\}$ e, na realidade, reduz-se a ela quando $m = 1$.

Para começar, consideremos a aplicação $f \colon \mathbb{R}^{m+1} - \{0\} \to S^m$ que chamaremos de *projeção radial* e definiremos por $f(x) = \dfrac{x}{|x|}\cdot$

Em cada ponto $x \neq 0$ de $\mathbb{R}^{m+1}$, a derivada de f é uma transformação linear $f'(x) \colon \mathbb{R}^{m+1} \to T_{f(x)}S^m$, cujo valor $f'(x) \cdot w$ num vetor arbitrário $w \in \mathbb{R}^{m+1}$ vamos determinar.

Se w é múltiplo de x então $f'(x) \cdot w = 0$ pois f é constante (igual a $x/|x|$) na semi-reta $0x$.

Como todo $w \in \mathbb{R}^{m+1}$ é soma, $w = c \cdot x + w_0$, de um múltiplo de x e um (único) vetor w_0, perpendicular a x (chamado a "projeção de w perpendicularmente a x"), vemos que $f'(x) \cdot w = f'(x) \cdot w_0$. Mas w_0, sendo perpendicular a x, é da forma $w_0 = \lambda'(0)$, onde λ é um caminho com $\lambda(0) = x$ e $|\lambda(t)|$ constante, igual a $|x|$. Então $(f \circ \lambda)(t) = \lambda(t)/|x|$, donde

$$f'(x) \cdot w = f'(x) \cdot w_0 = (f \circ \lambda)'(0) = \frac{\lambda'(0)}{|x|} = \frac{w_0}{|x|} = \frac{w - c \cdot x}{|x|}\cdot$$

Chamaremos *elemento de ângulo sólido* a forma diferencial $\alpha = f^*\omega$ em $\mathbb{R}^{m+1} - \{0\}$, induzida pelo elemento de volume ω da esfera S^m (orientada pela normal externa) através da projeção radial f.

Por definição, para $x \in \mathbb{R}^{m+1} - \{0\}$ e $w_1, \ldots, w_m \in \mathbb{R}^{m+1}$ arbitrários:

$$\begin{aligned}\alpha(x)(w_1, \ldots, w_m) &= \omega\left(\frac{x}{|x|}\right)(f'(x) \cdot w_1, \ldots, f'(x) \cdot w_m) = \\ &= \det.\left(\frac{x}{|x|}, \frac{w_1 - c_1 \cdot x}{|x|}, \ldots, \frac{w_m - c_m \cdot x}{|x|}\right) = \\ &= \frac{1}{|x|^{m+1}} \cdot \det(x, w_1, \ldots, w_m),\end{aligned}$$

pois o valor de um determinante não se altera quando se soma a uma de suas colunas um múltiplo de outra. Daí (como no Exemplo 16) vem:

$$\alpha(x) = \frac{1}{|x|^{m+1}} \sum_{i=1}^{m+1} (-1)^{i+1}\, x_i dx_1 \wedge \cdots \wedge \widehat{dx_i} \wedge \cdots \wedge dx_{m+1}\,.$$

Em particular, para $m = 1$, obtemos $\alpha(x,y) = \dfrac{xdy - ydx}{x^2 + y^2}$, que é a forma elemento de ângulo $\mathbb{R}^2 - \{0\}$, já conhecida desde o Capítulo 4.

Com exceção do fator $1/|x|^{m+1}$, a forma α coincide com a forma ω do Exemplo 16, a qual, restrita a S^m, dá o elemento de volume da esfera. Mas a semelhança acaba aí. As propriedades de α, como veremos no que se segue, são bem diferentes das de ω.

3 A diferencial exterior

Há três teoremas clássicos em Análise que generalizam o Teorema Fundamental do Cálculo, segundo o qual $\int_a^b df = f(b) - f(a)$. Eles são os seguintes:

Teorema de Green. *Seja $M \subset \mathbb{R}^2$ um compacto cujo bordo é uma curva ∂M. Se $a, b\colon M \to \mathbb{R}$ são funções de classe C^1 então vale a igualdade*

$$\iint_M \left(\frac{\partial b}{\partial x} - \frac{\partial a}{\partial y}\right) dxdy = \int_{\partial M} adx + bdy.$$

Teorema de Stokes. *Seja $M \subset \mathbb{R}^2$ uma superfície compacta cujo bordo é a curva ∂M. Se $a, b, c\colon M \to \mathbb{R}$ são funções de classe C^1 então*

$$\iint_M \left[\left(\frac{\partial b}{\partial x} - \frac{\partial a}{\partial y}\right) dxdy + \left(\frac{\partial c}{\partial y} - \frac{\partial b}{\partial z}\right) dydz + \left(\frac{\partial c}{\partial x} - \frac{\partial a}{\partial z}\right) dxdz\right] =$$
$$= \int_{\partial M} adx + bdy + cdz.$$

Teorema de Gauss. *Seja $M \subset \mathbb{R}^3$ uma região compacta cujo bordo é a superfície ∂M. Se $a, b, c \colon M \to \mathbb{R}$ são funções de classe C^1 então*

$$\iiint_M \left(\frac{\partial a}{\partial x} + \frac{\partial b}{\partial y} + \frac{\partial c}{\partial z}\right) dxdydz = \iint_{\partial M} adydz + bdzdx + cdxdy.$$

Esses três teoremas (cujos enunciados acima estão apenas esboçados) são casos particulares de um teorema geral, chamado também de Stokes, que demonstraremos no parágrafo 8. O enunciado do teorema geral é, concisamente, $\int_M d\omega = \int_{\partial M} \omega$. Nos 3 teoremas acima, se $\omega = adx + bdy$ então $d\omega = \left(\frac{\partial b}{\partial x} - \frac{\partial a}{\partial y}\right) dx \wedge dy$: se $\omega = adx + bdy + cdz$ então $d\omega = \left(\frac{\partial b}{\partial x} - \frac{\partial a}{\partial y}\right) dx \wedge dy + \left(\frac{\partial c}{\partial y} - \frac{\partial b}{\partial z}\right) dy \wedge dz + \left(\frac{\partial c}{\partial x} - \frac{\partial a}{\partial z}\right) dx \wedge dz$ e, finalmente, quando $\omega = ady \wedge dz + bdz \wedge dx + cdx \wedge dy$ então $d\omega = \left(\frac{\partial a}{\partial x} + \frac{\partial b}{\partial y} + \frac{\partial c}{\partial z}\right) dx \wedge dy \wedge dz$.

Neste parágrafo, definiremos o operador d, chamado a *diferenciação exterior*, o qual leva formas diferenciais de grau r em formas de grau $r+1$, e tem as expressões acima como casos particulares.

Começamos com formas em abertos do espaço euclidiano.

Seja $\omega = \Sigma\, a_I dx_I$ uma forma de grau r e classe C^k $(k \geq 1)$ no aberto $U \subset \mathbb{R}^m$. A *diferencial exterior* de ω é, por definição, a forma

$$d\omega = \sum_I da_I \wedge dx_I = \sum_{j,I} \frac{\partial a_I}{\partial x_j} dx_j \wedge dx_I\,.$$

Evidentemente, $d\omega$ é uma forma de grau $r+1$ e classe C^{k-1} em U.

Se ω é uma forma de grau zero, ou seja, uma função $\omega = a \colon U \to \mathbb{R}$, então $d\omega = da = \sum_j \frac{\partial a}{\partial x_j} dx_j\,.$

Se $\omega = \Sigma\, a_i dx_i$ é uma forma de grau 1 em U, então

$$\begin{aligned} d\omega &= \sum_i da_i \wedge dx_i = \sum_{i,j} \frac{\partial a_i}{\partial x_j}\, dx_j \wedge dx_i = \\ &= \sum_{j<i} \frac{\partial a_i}{\partial x_j}\, dx_j \wedge dx_i + \sum_{i<j} \frac{\partial a_i}{\partial x_j}\, dx_j \wedge dx_i = \\ &= \sum_{i<j} \frac{\partial a_j}{\partial x_i}\, dx_i \wedge dx_j - \sum_{i<j} \frac{\partial a_i}{\partial x_j}\, dx_i \wedge dx_j = \\ &= \sum_{i<j} \left(\frac{\partial a_j}{\partial x_i} - \frac{\partial a_i}{\partial x_j} \right) dx_i \wedge dx_j \,. \end{aligned}$$

[Na quarta igualdade, trocamos os nomes dos índices i, j no primeiro somatório e substituimos $dx_j \wedge dx_i$ por $-dx_i \wedge dx_j$ no segundo.]

Em particular:

Se $\omega = adx + bdy$ então $d\omega = \left(\dfrac{\partial b}{\partial x} - \dfrac{\partial a}{\partial y} \right) dx \wedge dy.$

Se $\omega = adx + bdy + cdz$, então

$$d\omega = \left(\frac{\partial b}{\partial x} - \frac{\partial a}{\partial y} \right) dx \wedge dy + \left(\frac{\partial c}{\partial y} - \frac{\partial b}{\partial z} \right) dy \wedge dz + \left(\frac{\partial c}{\partial x} - \frac{\partial a}{\partial z} \right) dx \wedge dz.$$

Seja agora $\omega = \sum\limits_{i=1}^{m} (-1)^{i+1}\, a_i dx_1 \wedge \cdots \wedge \widehat{dx_i} \wedge \cdots \wedge dx_m$ uma forma de grau $m-1$ no aberto $U \subset \mathbb{R}^m$. Então, como o produto exterior $dx_j \wedge dx_1 \wedge \cdots \wedge \widehat{dx_i} \wedge \cdots \wedge dx_m$ é nulo para $j \neq i$, temos

$$\begin{aligned} d\omega = \sum_{i=1}^{m} (-1)^{i+1} \frac{\partial a_i}{\partial x_i} dx_i \wedge dx_1 \wedge \cdots \wedge \widehat{dx_i} \wedge \cdots \wedge dx_m = \\ = \left(\sum_{i=1}^{m} \frac{\partial a_i}{\partial x_i} \right) dx_1 \wedge \cdots \wedge dx_m \,. \end{aligned}$$

Em particular, se $\omega = ady \wedge dz + bdz \wedge dx + cdx \wedge dy$ então $d\omega = \left(\dfrac{\partial a}{\partial x} + \dfrac{\partial b}{\partial y} + \dfrac{\partial c}{\partial z} \right) dx \wedge dy \wedge dz.$

Exemplo 18. $\omega = \dfrac{1}{m} \sum\limits_{i=1}^{m} (-1)^{i+1} x_i dx_1 \wedge \cdots \wedge \widehat{dx_i} \wedge \cdots \wedge dx_m$ é uma forma de grau $m-1$ em $\mathbb{R}^m$ cuja diferencial exterior é $d\omega = dx_1 \wedge \cdots \wedge dx_m$.

Em particular, para $m = 2$, temos $\omega = \frac{1}{2}(xdy - ydx)$. O fato de que $d\left[\frac{1}{2}(xdy - ydx)\right] = dx \wedge dy$ mostra, por meio do Teorema de Green, como calcular a área de uma região compacta $M \subset \mathbb{R}^2$, limitada por uma curva ∂M. Tem-se

$$\text{área de } M = \iint_M dxdy = \iint_M d\omega = \int_{\partial M} \omega = \frac{1}{2}\int_{\partial M}(xdy - ydx).$$

O cálculo da área de M reduz-se a uma integral curvilínea. A justificativa e o enunciado preciso dos princípios que intervêm neste exemplo serão fornecidos no parágrafo 8 adiante.

As principais propriedades da diferencial exterior são enunciadas a seguir.

Teorema 6. *Sejam $U \subset \mathbb{R}^m$, $V \subset \mathbb{R}^n$ abertos, $\omega, \bar{\omega}$ formas diferenciais de classe C^1 definidas em V e $f\colon U \to \mathbb{R}^n$ uma aplicação de classe C^2, tal que $f(U) \subset V$. Então:*

1. *Se $\omega\colon U \to \mathbb{R}$ é uma forma de grau zero (função real) então $d\omega$ é a diferencial usual de uma função;*
2. *$d(\omega + \bar{\omega}) = d\omega + d\bar{\omega}$;*
3. *Se ω é de classe C^2 então $dd\omega = 0$;*
4. *$d(\omega \wedge \bar{\omega}) = d\omega \wedge \bar{\omega} + (-1)^{gr\,\omega}\,\omega \wedge d\bar{\omega}$;*
5. *$d(f^*\omega) = f^*(d\omega)$.*

Demonstração: 1. e 2. são evidentes. Em virtude de 2., basta demonstrar as três propriedades seguintes no caso em que $\omega = adx_I$ e $\bar{\omega} = bdx_J$ são monômios. Vejamos 3. Se $\omega = adx_I$ então $d\omega = \sum_{j=1}^{n} \frac{\partial a}{\partial x_j} dx_j \wedge dx_I$ e daí

$$dd\omega = \left[\sum_{k,j=1}^{n} \frac{\partial^2 a}{\partial x_k \partial x_j} dx_k \wedge dx_j\right] \wedge dx_I =$$

$$= \left[\sum_{j<k} \left(\frac{\partial^2 a}{\partial x_j \partial x_k} - \frac{\partial^2 a}{\partial x_k \partial x_j}\right) dx_j \wedge dx_k\right] \wedge dx_I = 0,$$

em virtude do Teorema de Schwarz.

Quanto a 4. temos

$$\begin{aligned} d(\omega \wedge \bar{\omega}) = d(adx_I \wedge bdx_J) = d(abdx_I \wedge dx_J) = (bda + adb)\wedge \\ \wedge dx_I \wedge dx_J = bda \wedge dx_I \wedge dx_J + adb \wedge dx_I \wedge dx_J = \\ = (da \wedge dx_I) \wedge bdx_J + (-1)^{gr\,\omega}\, adx_I \wedge (db \wedge dx_J) = \\ = d\omega \wedge \bar{\omega} + (-1)^{gr\,\omega}\, \omega \wedge d\bar{\omega}. \end{aligned}$$

Finalmente, para provar 5., comecemos com o caso em que a forma ω se reduz a uma função $g\colon V \to \mathbb{R}$. Então, pela Regra da Cadeia, em todo ponto $x \in U$ tem-se $dg(f(x)) \cdot f'(x) = d(g \circ f)(x)$. Portanto, para qualquer $w \in \mathbb{R}^m$, vale

$$f^*(dg)(x) \cdot w = dg(f(x)) \cdot f'(x) \cdot w = d(g \circ f)(x) \cdot w.$$

Logo $f^*(dg) = d(g \circ f) = d(f^*g)$. Agora, consideremos uma forma $\omega = adx_I = adx_{i_1} \wedge \cdots \wedge dx_{i_r}$ de grau r arbitrário. Observemos preliminarmente que resulta de 4., 3. e uma indução óbvia que se $a\colon V \to \mathbb{R}$ é de classe C^1 e $g_1, \ldots, g_r\colon V \to \mathbb{R}$ são funções de classe C^2 então $d(adg_1 \wedge \cdots \wedge dg_r) = da \wedge dg_1 \wedge \cdots \wedge dg_r$. Lembremos ainda que $f^*(\alpha \wedge \beta) = f^*\alpha \wedge f^*\beta$. Então

$$f^*\omega = f^*a \cdot f^*dx_{i_1} \wedge \cdots \wedge f^*dx_{i_r} = f^*a \cdot d(x_{i_1} \circ f) \wedge \cdots \wedge d(x_{i_r} \circ f).$$

Logo:

$$\begin{aligned} df^*\omega = d(f^*a) \wedge d(x_{i_1} \circ f) \wedge \cdots \wedge d(x_{i_r} \circ f) = \\ = f^*da \wedge f^*dx_{i_1} \wedge \cdots \wedge f^*dx_{i_r} = \\ = f^*(da \wedge dx_{i_1} \wedge \cdots \wedge dx_{i_r}) = f^*(d\omega). \end{aligned}$$

A seguir, definiremos a diferencial exterior de uma forma ω numa superfície M. A forma ω precisa apenas de ser C^1, mas a superfície M será suposta pelo menos de classe C^2.

Seja então ω uma forma de grau r e classe C^1 em M. Para toda parametrização $\varphi\colon U_0 \to U$ em M, existe uma única forma $d_\varphi\omega$, de grau $r+1$ em U, tal que $\varphi^*(d_\varphi\omega) = d(\varphi^*\omega)$. Isto se deve ao fato de que o "pull back" $\omega \mapsto \varphi^*\omega$ é uma bijeção das formas em U sobre as formas em U_0. Se $\omega|U = \Sigma\, a_I(u)du_I$, então, como sabemos (Exemplo 12), $\varphi^*\omega = \Sigma\, a_I dx_I$, donde $d(\varphi^*\omega) = \Sigma\, da_I \wedge dx_I$. Logo, a igualdade $\varphi^*(d_\varphi\omega) = d(\varphi^*\omega)$ significa que $d_\varphi\omega = \Sigma\, da_I \wedge du_I$.

Vamos agora mostrar que se $\psi\colon V_0 \to V$ é outra parametrização então $d_\varphi\omega = d_\psi\omega$ em $U\cap V$. Com efeito, temos $\varphi = \psi\circ\xi\colon \varphi^{-1}(U\cap V) \to U\cap V$. Logo $\varphi^* = \xi^*\circ\psi^*$. Daí, em $U\cap V$:

$$\varphi^*(d_\psi\omega) = \xi^*\psi^*(d_\psi\omega) = \xi^* d(\psi^*\omega) = d(\xi^*\psi^*\omega) = \\ = d(\varphi^*\omega) = \varphi^*(d_\varphi\omega).$$

Portanto $d_\psi\omega = d_\varphi\omega$ em $U\cap V$. [Note que o uso do item 5. do Teorema 6 requer que as parametrizações φ, ψ sejam pelo menos de classe C^2.]

Definimos a *diferencial exterior* $d\omega$ da forma ω de grau r em M como a forma de grau $r+1$ em M cujo valor em cada ponto $x\in M$ é dado por $(d\omega)(x) = (d_\varphi\omega)(x)$, onde $\varphi\colon U_0\to U$ é qualquer parametrização em M tal que $x\in U$.

Como vimos acima, temos $(d\omega)(x) = \Sigma\, da_I\wedge du_I$ para todo $x = \varphi(u)\in U$, se $\omega(x) = \Sigma\, a_I du_I$ for a expressão de ω relativa à parametrização φ. Vimos também que a definição acima produz uma forma diferencial bem determinada $d\omega$, independente das parametrizações utilizadas. Em particular se $\omega = \Sigma\, a_I du_I = \Sigma\, b_I dv_I$ então $\Sigma\, da_I\wedge du_I = \Sigma\, db_I\wedge dv_I = d\omega$.

Se ω é de classe C^k, a expressão $d\omega = \Sigma\, da_I\wedge du_I$ mostra que $d\omega$ é de classe C^{k-1}. Além disso, valem para a diferencial exterior de formas em superfícies as propriedades expressas nos itens 1. a 5. do Teorema 6, sendo que, no item 5., tem-se uma aplicação, de classe C^2, $f\colon M\to N$. Como essas propriedades são todas locais, as demonstrações consistem simplesmente em tomar parametrizações e fazer o caso recair no Teorema 6.

Seja ω uma forma diferencial de grau r e classe C^1 numa superfície M. Diz-se que ω é *fechada* quando $d\omega = 0$ e que ω é *exata* quando existe uma forma α em M, de grau $r-1$ e classe C^2, tal que $d\alpha = \omega$.

Toda forma exata é fechada, pois $dd\alpha = 0$. A recíproca é falsa como mostra o elemento de ângulo em $\mathbb{R}^2 - \{0\}$. (Evidentemente, as definições que acabamos de dar são compatíveis com as que foram dadas para formas de grau 1, no Capítulo IV.)

Diz-se que ω é uma forma de *grau máximo* quando tem grau m e está definida numa superfície de dimensão m. Toda forma ω de classe C^1 e grau máximo é fechada pois $d\omega$ tem grau $m+1$ e toda forma $(m+1)$-linear alternada num espaço vetorial de dimensão m é nula.

Seja $f\colon M\to N$ uma aplicação de classe C^2. A igualdade $f^*(d\omega) = d(f^*\omega)$ mostra que se ω é uma forma fechada (respect. exata) em N

então $f^*\omega$ é fechada (respect. exata) em M. Isto se aplica, em particular, à restrição de uma forma a uma superfície contida em seu domínio.

Segue-se da observação acima que se ω é uma forma de grau máximo em N então $f^*\omega$ é uma forma fechada em M (mesmo que não seja de grau máximo). Daí concluimos, sem fazer cálculos, que o elemento de ângulo sólido é uma forma fechada de grau m em $\mathbb{R}^{m+1} - \{0\}$, pois é induzida por uma forma de grau máximo em S^m, a saber, o elemento de volume. Resultará do Teoema de Stokes que o elemento de ângulo sólido não é uma forma exata em $\mathbb{R}^{m+1} - \{0\}$.

Um exemplo de forma exata é o elemento de volume $\omega = dx_1 \wedge \cdots \wedge dx_m$ em $\mathbb{R}^m$ (e portanto em qualquer aberto $U \subset \mathbb{R}^m$). Com efeito, sabemos que $\omega = d\beta$, onde $\beta = \dfrac{1}{m} \sum\limits_{i=1}^{m} (-1)^{i+1}\, x_i dx_1 \wedge \cdots \wedge \widehat{dx_i} \wedge \cdots \wedge dx_m$. Uma explicação geral para isto é a seguinte: ω é fechada porque tem grau máximo e, pelo Exercício 3.4, adiante, toda forma fechada cujo domínio é todo o espaço $\mathbb{R}^m$ é exata.

Nos Exercícios 3.6 a 3.10 e 8.20 deste capítulo, são estudadas as relações entre formas fechadas e exatas numa superfície, de maneira semelhante ao que fizemos no Capítulo IV para abertos de $\mathbb{R}^2$. Um instrumento importante nesse estudo é a integral de superfície que definiremos no parágrafo 6.

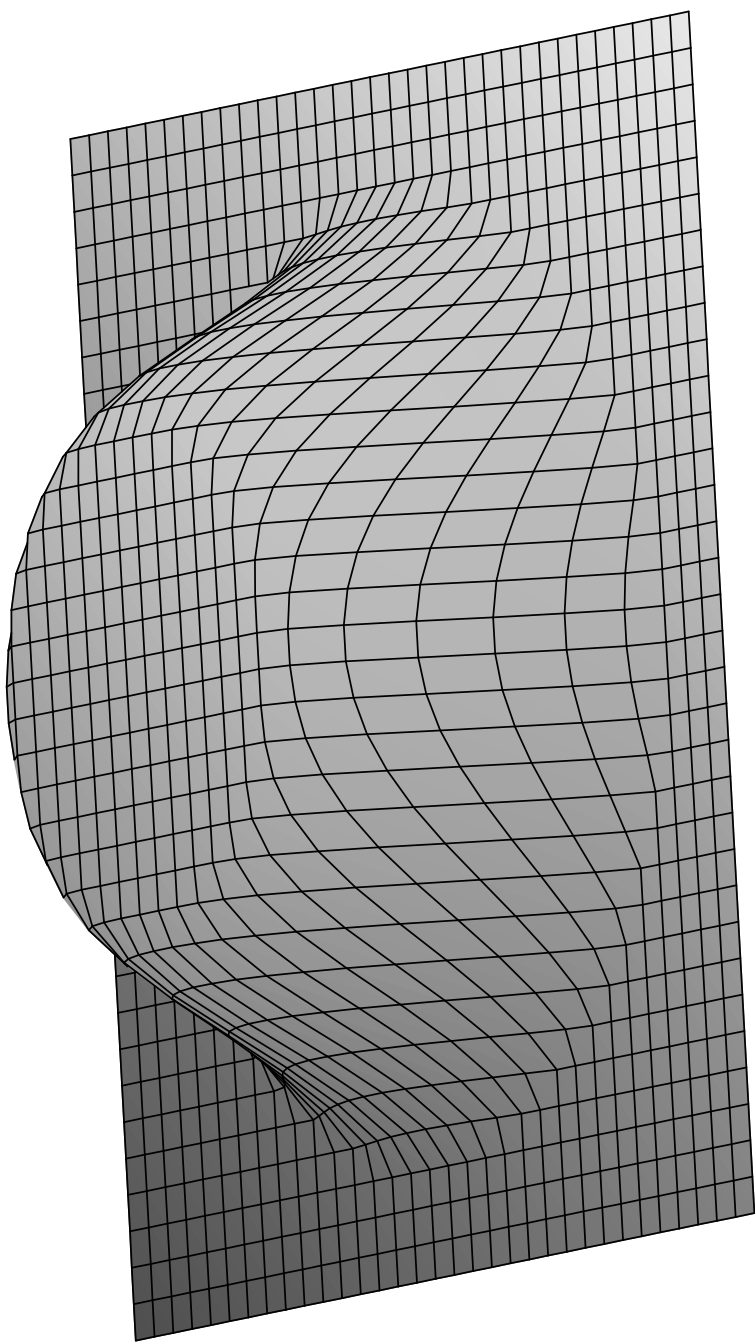
4 Partições da unidade

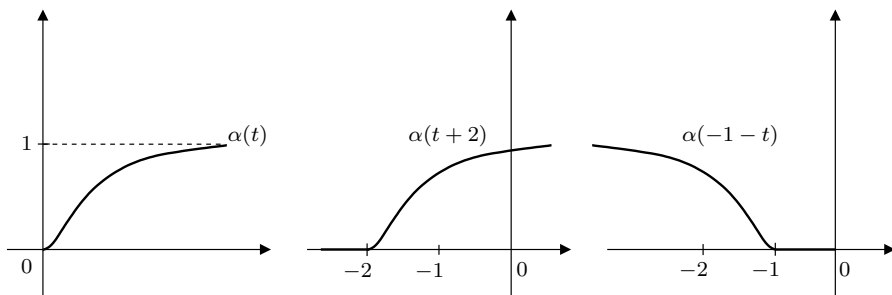
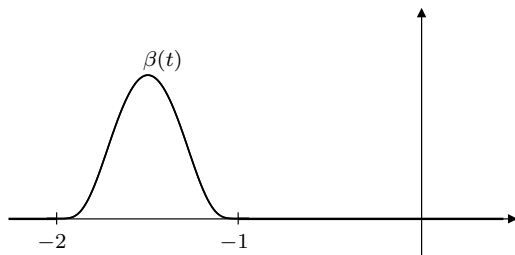
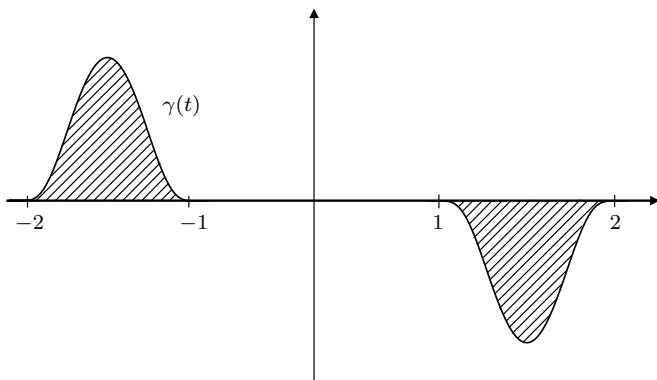
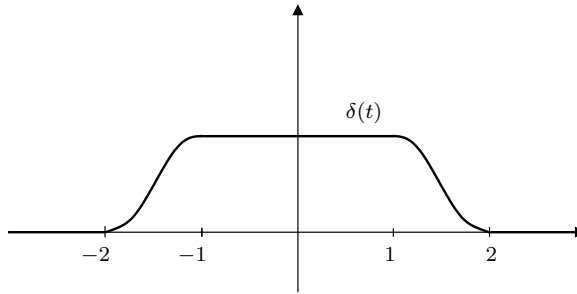
No parágrafo 3, mostramos como se define a integral, sobre uma superfície m-dimensional orientada, de uma forma diferencial contínua de grau m, cujo suporte é compacto e está contido numa vizinhança parametrizada. Esta última restrição será eliminada se mostrarmos que toda forma diferencial em questão pode ser escrita como uma soma $\omega = \Sigma\, \omega_i$, onde cada ω_i é contínua e tem suporte compacto, contido numa vizinhança parametrizada. Poremos, então, $\int_M \omega = \Sigma \int_M \omega_i$. A decomposição $\omega = \Sigma\, \omega_i$ será feita mediante a técnica das partições da unidade, que desenvolveremos neste parágrafo.

A maneira clássica de definir a integral de superfície $\int_M \omega$ (que pode ser encontrada em quase todos os livros de Cálculo), consiste em decompor a superfície como reunião $M = \cup K_i$ de conjuntos compactos J-mensuráveis sem pontos interiores em comum, cada um deles contido numa vizinhança parametrizada, e pôr $\int_M \omega = \Sigma \int_{K_i} \omega$. Isto pode ser

feito rigorosamente, mas a tarefa de provar que existe a decomposição $M = \cup K_i$, e que a soma $\Sigma \int_{K_i} \omega$ não depende dela, é bastante laboriosa e certamente se situa fora do nível e do âmbito de interesse deste livro. É mais simples decompor o integrando ω. Além disso, o método das partições da unidade, que exporemos agora, tem várias outras aplicações.

Indiquemos com $B(r)$ a bola aberta de centro 0 e raio r em $\mathbb{R}^m$ e com $B[r]$ a bola fechada correspondente.

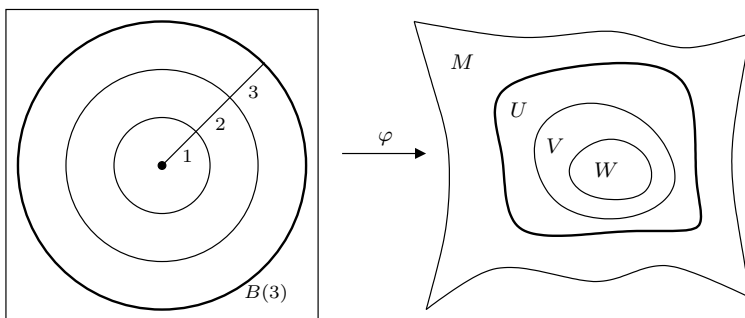
Nosso próximo passo será definir uma função $\xi\colon \mathbb{R}^m \to \mathbb{R}$, de classe C^∞, tal que $0 \leq \xi(x) \leq 1$ para todo x, $\xi|B[1] = 1$ e $\xi(x) = 0$ se $x \notin B[2]$. Quando $1 < |x| < 2$, tem-se $0 < \xi(x) < 1$. O gráfico de ξ, no caso $m = 2$, tem a forma de uma bacia emborcada.

A função ξ pode ser definida mais diretamente. Entretanto, acha-

mos instrutivo fazê-lo por etapas. Começamos com a clássica função de Cauchy, $\alpha\colon \mathbb{R} \to \mathbb{R}$, $\alpha(t) = 0$ se $t \leq 0$ e $\alpha(t) = \exp(-1/t)$ se $t > 0$. Abaixo, vemos os gráficos de $\alpha(t)$, $\alpha(t+2)$ e $\alpha(-1-t)$.

Em seguida, definimos $\beta\colon \mathbb{R} \to \mathbb{R}$ pondo $\beta(t) = \alpha(t+2) \cdot \alpha(-1-t)$ para todo $t \in \mathbb{R}$. Sabemos que $\alpha \in C^\infty$, logo $\beta \in C^\infty$. O gráfico de β é:

Seja $b = \int_{-\infty}^{+\infty} \beta(t)dt = \int_{-2}^{-1} \beta(t)\,dt$. Então $\dfrac{1}{b}\int_{-2}^{-1} \beta(t)dt = 1$. Definamos agora uma nova função, $\gamma\colon \mathbb{R} \to \mathbb{R}$, de classe C^∞, pondo $\gamma(t) = \beta(t)/b$ se $t \leq 0$ e $\gamma(t) = -\beta(-t)/b$ se $t > 0$. Segue-se o gráfico de γ. As regiões hachuradas têm área 1.

A próxima função, $\delta\colon \mathbb{R} \to \mathbb{R}$, também de classe C^∞, é definida por $\delta(t) = \int_{-\infty}^{t} \gamma(s)\, ds = \int_{-2}^{t} \gamma(s) ds$. Seu gráfico é:

Finalmente, a função $\xi\colon \mathbb{R}^m \to \mathbb{R}$ é dada por $\xi(x) = \delta(|x|)$, onde $|x| = \sqrt{\Sigma x_i^2}$. Como ξ é constante na bola $B[1]$, a não-diferenciabilidade da raiz quadrada no ponto 0 não impede que se tenha $\xi \in C^\infty$.

Para todo ponto p numa superfície M, de classe C^k, existem funções de classe C^k, definidas em M, análogas a ξ, com suportes em vizinhanças arbitrariamente pequenas de p. Isto se segue do

Lema 1. *Seja $p \in A \subset M$, onde A é um aberto na superfície M, de classe C^k. Existe uma parametrização $\varphi\colon B(3) \to U$ em M, com $\varphi \in C^k$, $\varphi(0) = p$ e $U \subset A$.*

Demonstração: Seja ψ uma parametrização de classe C^k em M, definida num aberto $U_0 \subset \mathbb{R}^m$, o qual contém um ponto a tal que $\psi(a) = p$. Pela continuidade de ψ, existe $r > 0$ tal que $\psi(B(a;r)) \subset A$. Seja h o difeomorfismo C^∞ de $\mathbb{R}^m$ sobre si mesmo, definido por $h(x) = a + \frac{r}{3} \cdot x$. Então $h(B(3)) = B(a;r)$ e $h(0) = a$, de modo que, pondo $U = \psi(B(a;r))$, obtemos uma parametrização $\varphi = \psi \circ h\colon B(3) \to U$, como o lema requer.

Quando tivermos uma parametrização φ do tipo acima, usaremos as notações $U = \varphi(B(3))$, $V = \varphi(B(2))$ e $W = \varphi(B(1))$. Tem-se $p \in W \subset V \subset U$. Evidentemente, $\overline{W}$ e $\overline{V}$ são compactos. A demonstração do Lema 1 mostra que, se for preciso, podemos supor também que $\overline{U}$ é compacto e difeomorfo a $B[3]$.

Nessas condições, definimos a função $\xi_\varphi\colon M \to \mathbb{R}$, de classe C^k, pondo $\xi_\varphi(x) = \xi[\varphi^{-1}(x)]$ se $x \in U$ e $\xi_\varphi(x) = 0$ se $x \in M - V$. Vemos que ξ_φ é uma função de classe C^k, igual a 1 em todos os pontos de $\overline{W}$ e nula em $M - V$. Na realidade, temos $\operatorname{supp}(\xi_\varphi) = \overline{V}$. Na faixa $V - \overline{W}$, temos $0 < \xi_\varphi(x) < 1$. Diremos que ξ_φ é a *função auxiliar* associada à parametrização φ.

O próximo ingrediente necessário para o estudo das partições da unidade é a noção de família localmente finita.

Uma família $\mathcal{C} = (C_\lambda)_{\lambda \in L}$ de subconjuntos de uma superfície M diz-se *localmente finita* quando todo ponto $p \in M$ possui uma vizinhança V em M que intersecta apenas um número finito de conjuntos C_λ. Mais precisamente: dado p, existem V e um subconjunto finito $L_0 \subset L$ tais que $\lambda \notin L_0 \Rightarrow V \cap C_\lambda = \varnothing$. Quando, além disso, se tem $M = \bigcup_{\lambda \in L} C_\lambda$, diz-se que $\mathcal{C} = (C_\lambda)_{\lambda \in L}$ é uma *cobertura localmente finita* de M.

Exemplo 19. Toda família finita é localmente finita. Se $\mathcal{C} = (C_\lambda)_{\lambda \in L}$ é uma cobertura aberta de M, tal que cada conjunto C_λ tem interseção não vazia com apenas um número finito de outros conjuntos da família então $\mathcal{C}$ é localmente finita: para cada $p \in M$, tome um C_λ contendo p e ponha $V = C_\lambda$. Caso particular: $\mathcal{C} = (C_i)_i \in \mathbb{N}$, onde $C_i = \{x \in \mathbb{R}^m; i - 1 < |x| < i + 1\}$. Mais particular ainda: para $m = 1$, $C_i = (i - 1, i + 1)$. Então C_i intersecta apenas C_{i-1} e C_{i+1}. Convém notar que nem toda família localmente finita é desse tipo. Por exemplo, sejam $C_1 = \mathbb{R}^m$ e, para cada $i \in \mathbb{N}$, $C_{i+1} = \mathbb{R}^m - B[i]$, complementar da bola de centro 0 e raio i em $\mathbb{R}^m$. A família $\mathcal{C} = (C_i)_{i \in \mathbb{N}}$ é uma cobertura localmente finita de $\mathbb{R}^m$ mas cada conjunto C_i intersecta *todos* os demais. [Para ver que $\mathcal{C}$ é localmente finita, dado $p \in \mathbb{R}^m$ tome $j \in \mathbb{N}$ tal que $|p| < j$. Então a vizinhança $V = B(j)$ do ponto p intersecta apenas os conjuntos C_i com $i < j$.] Finalmente, $(B(i))_{i \in \mathbb{N}}$ não é uma família localmente finita.

Exemplo 20. Uma família $\mathcal{C} = (C_\lambda)_{\lambda\in L}$ de subconjuntos de M diz-se *pontualmente finita* quando cada ponto $p \in M$ pertence apenas a um número finito de conjuntos C_λ. Toda família localmente finita é pontualmente finita. A recíproca é falsa: cada ponto $p \in \mathbb{R}$ pertence no máximo a um número finito de intervalos $(1/n, 2/n)$, $n \in \mathbb{N}$, mas toda vizinhança de 0 em $\mathbb{R}$ intersecta uma infinidade de tais intervalos.

Nosso interese em famílias localmente finitas resulta da seguinte observação.

Seja $(\omega_\lambda)_{\lambda\in L}$ uma família de formas diferenciais de grau r e classe C^k numa superfície M. Se a família $(\text{supp.}\,\omega_\lambda)_{\lambda\in L}$ dos seus suportes é pontualmente finita, então a soma

$$\omega = \sum_{\lambda\in L} \omega_\lambda$$

tem sentido e é uma forma de grau r em M. De fato, para cada $x \in M$ existe um conjunto finito de índices $L_0 = \{\lambda_1, \ldots, \lambda_s\} \subset L$ tal que $\omega_\lambda(x) = 0$ se $\lambda \notin L_0$. Pomos então, por definição

$$\omega(x) = \omega_{\lambda_1}(x) + \cdots + \omega_{\lambda_s}(x).$$

Na definição acima, para cada $x \in M$ o conjunto L_0 varia, de modo que não se pode garantir que $\omega \in C^k$. A hipótese de que os suportes das ω_λ formam uma família pontualmente finita não basta.

Se, entretanto, a família $(\text{supp.}\,\omega_\lambda)_{\lambda\in L}$ é *localmente finita* e cada $\omega_\lambda \in C^k$ então $\omega = \Sigma\,\omega_\lambda \in C^k$. Com efeito, neste caso, para cada $p \in M$ existem uma vizinhança V de p em M e índices $\lambda_1, \ldots, \lambda_s \in L$ tais que

$$\omega(x) = \omega_{\lambda_1}(x) + \cdots + \omega_{\lambda_s}(x)$$

para todo $x \in V$.

A seguir, registraremos três propriedades das famílias localmente finitas:

Propriedade 1. *Toda família localmente finita numa superfície M é enumerável. Mais precisamente: se $\mathcal{C} = (C_\lambda)_{\lambda\in L}$ é localmente finita então existe um subconjunto enumerável $L_* \subset L$ tal que $\lambda \notin L_* \Rightarrow C_\lambda = \varnothing$.*

Demonstração: Para todo ponto $p \in M$ existe um aberto $V_p \subset M$ que o contém, tal que V_p intersecta apenas um número finito de conjuntos C_λ. Da cobertura aberta $M = \cup V_p$ podemos, pelo Teorema de Lindelöf,

escolher uma subcobertura enumerável, digamos $M = \bigcup_{i=1}^{\infty} V_i$. Para cada $i \in \mathbb{N}$, seja L_i o conjunto (finito) dos índices λ tais que $C_\lambda \cap V_i \neq \varnothing$. Pondo $L_* = \bigcup_{i=1}^{\infty} L_i$, vemos que L_* é enumerável e que $\lambda \notin L_* \Rightarrow C_\lambda \cap V_i = \varnothing$ para todo $i \in \mathbb{N}$, donde $C_\lambda = C_\lambda \cap M = \bigcup_{i=1}^{\infty} (C_\lambda \cap V_i) = \varnothing$.

Sempre que for conveniente, podemos então escrever uma família localmente finita como $\mathcal{C} = (C_i)_{i \in \mathbb{N}}$.

Propriedade 2. *Se $C = (C_\lambda)_{\lambda \in L}$ é uma família localmente finita na superfície M, um conjunto compacto $K \subset M$ só intersecta no máximo um número finito de conjuntos C_λ. Em particular, se M é compacta, toda família localmente finita de subconjuntos de M é finita. (Isto é, $C_\lambda = \varnothing$ salvo um número finito de índices λ.)*

Demonstração: Para todo ponto $p \in K$, existe V_p aberto em M, com $p \in V_p$, tal que cada V_p intersecta apenas um número finito de conjuntos C_λ. Da cobertura $K \subset \cup V_p$ escolhemos, pela compacidade de K, uma subcobertura finita $K \subset V_1 \cup \cdots \cup V_s$. Para cada $i = 1, \ldots, s$, seja L_i o conjunto (finito) dos índices λ tais que $C_\lambda \cap V_i \neq \varnothing$. Pondo $L_0 = \bigcup_{i=1}^{s} L_i$, vemos que L_0 é finito e que $\lambda \notin L_0 \Leftrightarrow C_\lambda \cap K \subset \bigcup_{i=1}^{s} (C_\lambda \cap V_i) = \varnothing$.

Propriedade 3. *Se a família $\mathcal{C} = (C_\lambda)_{\lambda \in L}$ é localmente finita então $\overline{\bigcup_{\lambda \in L} C_\lambda} = \bigcup_{\lambda \in L} \overline{C}_\lambda$.*

Demonstração: Estamos considerando o fecho *relativamente* à superfície M: se $X \subset M$, então $\overline{X} = \{x \in M; d(x, X) = 0\}$. Sem nenhuma hipótese sobre os C_λ, tem-se sempre a inclusão $\cup \overline{C}_\lambda \subset \overline{\cup C_\lambda}$. No caso localmente finito, vale a inclusão oposta. Com efeito, seja $p \in M$ tal que $p \notin \cup \overline{C}_\lambda$. Existe $r > 0$ tal que, pondo $V = \{x \in M; |x - p| < r\}$ então V intersecta no máximo os conjuntos $C_{\lambda_1}, \ldots C_{\lambda_s}$ da família $\mathcal{C}$. Isto quer dizer que $\lambda \notin \{\lambda_1, \ldots, \lambda_s\} \Rightarrow d(p, C_\lambda) \geq r$. Além disso, como $p \notin \overline{C}_\lambda$ para cada λ, temos $d_i = d(p, C_{\lambda_i}) > 0$ para $i = 1, \ldots, s$. Seja d o menor dos números positivos $d_1, \ldots, d_s, r$. Para todo $x \in \cup C_\lambda$ temos $|x - p| \geq d$. Assim $d(p, \cup C_\lambda) \geq d > 0$, e daí $p \notin \overline{\cup C_\lambda}$. Provamos que $p \notin \cup \overline{C}_\lambda \Rightarrow p \notin \overline{\cup C_\lambda}$, o que dá a inclusão que procurávamos.

Corolário. *Se $(F_\lambda)_{\lambda \in L}$ é uma família localmente finita de subconjuntos fechados de M então $F = \bigcup_{\lambda \in L} F_\lambda$, é um subconjunto fechado em M.*

Seja M uma superfície de classe C^k. Uma *partição da unidade* de classe C^k em M é uma família de funções $(\xi_\lambda)_{\lambda \in L}$ de classe C^k em M, tais que:

1) Para quaisquer $\lambda \in L$ e $x \in M$, tem-se $\xi_\lambda(x) \geq 0$;

2) A família $\mathcal{C} = (\text{supp.}\, \xi_\lambda)_{\lambda \in L}$ é localmente finita em M;

3) Para todo $x \in M$, tem-se $\sum\limits_{\lambda \in L} \xi_\lambda(x) = 1$.

Em virtude da condição 2), a soma em 3) é finita em cada ponto $x \in M$ e, na realidade, localmente finita. Além disso, tem-se $0 \leq \xi_\lambda(x) \leq 1$ para $x \in M$ e $\lambda \in L$ quaisquer, por causa de 1) e 3). Note ainda que, para todo $x \in M$ existe pelo menos um índice λ tal que $\xi_\lambda(x) > 0$, logo os suportes das funções ξ_λ constituem uma cobertura da superfície M.

A definição acima inclui o caso de partições finitas da unidade: $\xi_1 + \cdots + \xi_s = 1$. Isto ocorre quando $\xi_\lambda \equiv 0$ salvo para um número finito de índices λ. *Numa superfície compacta, toda partição da unidade é finita*, devido à Propriedade 2 acima. [Dizer que uma função tem suporte vazio significa afirmar que ela é identicamente nula.]

Diz-se que a partição da unidade $\sum\limits_{\alpha \in A} \xi_\alpha = 1$ está *subordinada* à cobertura $\mathcal{C} = (C_\lambda)_{\lambda \in L}$ da superfície M quando, para cada $\alpha \in A$, existe algum $\lambda \in L$ tal que $\text{supp}\,.\, \xi_\alpha \subset C_\lambda$.

Intuitivamente, a cobertura $\mathcal{C}$ indica um grau de pequenez (ou "fineza") dos suportes das funções ξ_α , no seguinte sentido:

Dadas duas coberturas $\mathcal{C}$, $\mathcal{C}'$ numa superfície M, dizemos que $\mathcal{C}$ é *mais fina* que $\mathcal{C}'$, ou que $\mathcal{C}$ *refina* $\mathcal{C}'$ ou $\mathcal{C}$ é um *refinamento* de $\mathcal{C}'$, quando, para todo $C \in \mathcal{C}$ existe algum $C' \in \mathcal{C}'$ tal que $C \subset C'$.

Nestes termos, a partição da unidade $\Sigma\, \xi_\alpha = 1$ está subordinada à cobertura $\mathcal{C} = (C_\lambda)$ quando os suportes das funções ξ_α formam uma cobertura que refina $\mathcal{C}$.

Dizemos que a partição da unidade $\sum\limits_{\lambda \in L} \xi_\lambda = 1$ é *estritamente subordinada* à cobertura $\mathcal{C}$ quando $\mathcal{C} = (C_\lambda)_{\lambda \in L}$ tem índices no mesmo conjunto que as funções ξ_λ e, além disso, $\text{supp}(\xi_\lambda) \subset C_\lambda$ para todo $\lambda \in L$.

Estamos agora em condições de enunciar o principal resultado deste parágrafo.

Teorema 7. *Seja M uma superfície de classe C^k. Para toda cobertura aberta $\mathcal{C} = (C_\lambda)_{\lambda \in L}$ de M existe uma partição da unidade $\sum\limits_{\lambda \in L} \xi_\lambda = 1$, de classe C^k, estritamente subordinada à cobertura $\mathcal{C}$.*

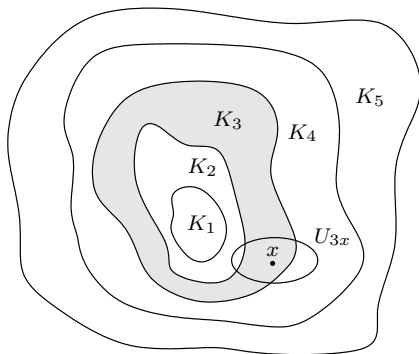
A demonstração se baseia nos dois lemas seguintes.

Lema 2. *Toda superfície M se escreve como reunião enumerável $M = \cup K_i$ de compactos tais que $K_i \subset \mathrm{int}\,.K_{i+1}$ para todo $i \in \mathbb{N}$.*

Demonstração: Cada ponto $x \in M$ está contido num aberto $V_x = \varphi(B(2))$, como no Lema 1, onde $\overline{V}_x$ é compacto. Da cobertura $M = \bigcup_{x \in M} V_x$ extraimos, pelo Teorema de Lindelöf, uma subcobertura enumerável $M = \cup V_i$. Pondo $L_i = \overline{V}_i$, cada L_i é compacto e vale ainda $M = \cup L_i$. Definimos os K_i por indução. Pomos $K_1 = L_1$ e, supondo obtidos $K_1, \ldots, K_i$ de modo que $K_j \subset \mathrm{int}\, K_{j+1}$ para $j = 1, \ldots, i-1$ e $K_i \supset L_1 \cup \cdots \cup L_i$, recobrimos o compacto $K_i \cup L_{i+1}$ com um número finito de conjuntos V_j e chamamos de K_{i+1} a reunião dos L_j correspondentes.

Observação: Se a superfície M é compacta, o lema acima é trivial. Podemos tomar $K_i = M$ para todo i. Ou então, seguindo a demonstração acima, notar que, a partir de um certo ponto $K_i = K_{i+1} = \cdots = M$.

Lema 3. *Toda cobertura aberta $\mathcal{C}$ de uma superfície M de classe C^k pode ser refinada por uma cobertura aberta localmente finita, formada por imagens de parametrizações $\varphi \colon B(3) \to U$, de classe C^k em M, tais que os abertos $W = \varphi(B(1))$ ainda cobrem M.*

Demonstração: Usando o Lema 2, escrevemos $M = \cup K_i$, onde cada K_i é compacto e $K_i \subset \mathrm{int}\,.K_{i+1}$ para todo $i \in \mathbb{N}$. Todo ponto $x \in K_2$ pertence a algum $C \in \mathcal{C}$. Aplicando o Lema 1, com $A = C \cap \mathrm{int}\,.K_3$, concluimos que todo x em K_2 pertence a um conjunto W_{2x} tal que o U_{2x} correspondente está contido em $\mathrm{int}\,.K_3$ e em algum aberto de $\mathcal{C}$. Da cobertura $K_2 \subset \bigcup_x W_{2x}$ extraimos uma subcobertura finita. Seja $\mathcal{C}_2$ a cobertura finita de K_2 formada pelos conjuntos U_{2x} correspondentes. Analogamente, cada ponto x da "faixa" compacta $K_3 - \mathrm{int}\,.K_2$ pertence a algum conjunto W_{3x} (na forma do Lema 1) tal que o U_{3x} correspondente está contido em $\mathrm{int}\,.K_4$, em algum conjunto de $\mathcal{C}$ e é disjunto de K_1.

Da cobertura $K_3 - \operatorname{int}.K_2 \subset \bigcup_x W_{3x}$ extraimos uma subcobertura finita. Seja $\mathcal{C}_3$ a cobertura finita de $K_3 - \operatorname{int}.K_2$ formada pelos conjuntos U_{3x} correspondentes. Prosseguindo analogamente, obtemos, para cada $i \geq 3$, uma cobertura finita $\mathcal{C}_i$ da faixa compacta $K_i - \operatorname{int}.K_{i-1}$ por abertos do tipo U, cada um deles contido em $\operatorname{int}.K_{i+1}$, em algum conjunto de $\mathcal{C}$ e disjunto de K_{i-2}, de modo que os W correspondentes ainda cobrem a faixa. A reunião $\mathcal{C}' = \mathcal{C}_2 \cup \cdots \cup \mathcal{C}_i \cup \ldots$ é uma cobertura aberta de M por conjuntos do tipo U, tal que os W correspondentes ainda cobrem M. Cada $U \in \mathcal{C}'$ pertence a algum $\mathcal{C}_i$, logo intersecta, no máximo, um número finito de outros conjuntos de $\mathcal{C}'$, a saber, os pertencentes a $\mathcal{C}_{i-1} \cup \cdots \cup \mathcal{C}_{i+2}$. Portanto $\mathcal{C}'$ é um refinamento localmente finito de $\mathcal{C}$.

Demonstração do Teorema 7. Pelo Lema 3, existe um refinamento localmente finito $\mathcal{C}' = (U_i)_{i\in\mathbb{N}}$ da cobertura dada $\mathcal{C}$, com $U_i = \varphi_i(B(3))$. Para cada $i \in \mathbb{N}$, seja $\xi_i^*\colon M \to \mathbb{R}$ a função auxiliar de classe C^k associada à parametrização φ_i. Os suportes $\overline{V}_i = \operatorname{supp}.\xi_i^*$ formam uma cobertura localmente finita de M, que refina $\mathcal{C}$. Logo $\xi^* = \Sigma\,\xi_i^*$ é uma função de classe C^k, positiva em todos os pontos de M. As funções $\eta_i\colon M \to \mathbb{R}$, definidas por $\eta_i = \xi_i^*/\xi^*$, cumprem $\Sigma\,\eta_i = 1$, $\operatorname{supp}.\eta_i = \overline{V}_i$ e constituem uma partição da unidade de classe C^k, subordinada a $\mathcal{C}$. Para obter uma partição estritamente subordinada a $\mathcal{C}$, comecemos definindo uma "função escolha" $f\colon \mathbb{N} \to L$, isto é, para cada $i \in \mathbb{N}$ escolhamos um índice $\lambda = f(i) \in L$ tal que $\operatorname{supp}\eta_i = \overline{V}_i \subset C_\lambda$. Para cada $\lambda \in L$, ponhamos $\xi_\lambda = \sum_{f(i)=\lambda} \eta_i$. Com os $\overline{V}_i$ formam uma família

localmente finita, temos

$$\text{supp}\,.\,\xi_\lambda = \overline{\bigcup_{f(i)=\lambda} V_i} = \bigcup_{f(i)=\lambda} \overline{V}_i \subset C_\lambda\,.$$

A família $(\text{supp}\,.\,\xi_\lambda)_{\lambda\in L}$ é localmente finita. Com efeito, cada ponto $p \in M$ tem uma vizinhança V_p que intersecta no máximo $\overline{V}_{i_1}, \ldots, \overline{V}_{i_j}$, logo V_p intersecta $\text{supp}\,.\,\xi_\lambda$ somente quando $\lambda = f(i_1), \ldots,$ ou $\lambda = f(i_j)$. Então $\sum_{\lambda\in L} \xi_\lambda = 1$ é uma partição da unidade estritamente subordinada à cobertura $\mathcal{C} = (C_\lambda)_{\lambda\in L}$.

5 Aplicações da partição da unidade

Nos itens abaixo, mostraremos como se pode usar a técnica das partições da unidade; no primeiro exemplo, como método de passagem do local ao global; no segundo, como um processo de obter médias ponderadas. O terceiro é uma aplicação à Topologia. No parágrafo seguinte, a partição da unidade será empregada numa redução do global ao local, a fim de definir a integral de uma forma com suporte compacto numa superfície orientável.

Item A. *Aplicações diferenciáveis em conjuntos arbitrários.*

Seja $X \subset \mathbb{R}^m$. Dizemos que uma aplicação $f\colon X \to \mathbb{R}^n$ *estende-se localmente em classe* C^k $(k \geq 0)$ quando, para cada ponto $p \in X$, existe uma aplicação $F_p\colon Z_p \to \mathbb{R}^n$, de classe C^k numa vizinhança aberta Z_p de p em $\mathbb{R}^m$, tal que $F_p(x) = f(x)$ para todo $x \in Z_p \cap X$. Noutras palavras, f se estende, numa vizinhança de cada ponto de X, a uma aplicação de classe C^k. Se $f\colon X \to \mathbb{R}^n$ estende-se localmente em classe C^k então f é contínua.

Exemplo 21. Seja $X = \{(x,y) \in \mathbb{R}^2; x \neq 0\} \cup \{(0,0)\}$. A função contínua $f\colon X \to \mathbb{R}$, definida por $f(x,y) = y \cdot \text{sen}\,\dfrac{1}{x}$ se $x \neq 0$ e $f(0,0) = 0$, não se estende localmente em classe C^0, pois não admite extensão contínua a vizinhança alguma da origem.

Teorema 8. *Seja $M \subset \mathbb{R}^{m+r}$ uma superfície de dimensão m e classe C^q $(q \geq 1)$. Toda aplicação $f\colon M \to \mathbb{R}^n$ de classe C^k $(0 \leq k \leq q)$ estende-se localmente em classe C^k.*

A demonstração se baseia no lema abaixo.

Uma aplicação $\rho\colon U \to S$, de classe C^k no aberto $U \subset \mathbb{R}^m$, chama-se uma *retração* quando $S \subset U$ e $\rho(x) = x$ para todo $x \in S$. Tem-se então $S = \rho(U)$ e $\rho \circ \rho = \rho$. Diz-se, neste caso, que S é um *retrato de classe* C^k do aberto U.

Lema. *Sejam $M \subset \mathbb{R}^{m+r}$ uma superfície de dimensão m e classe C^k $(k \geq 1)$. Todo ponto $p \in M$ está contido num aberto Z de $\mathbb{R}^{m+r}$ tal que $V = Z \cap M$ é um retrato de classe C^k de Z.*

Demonstração: Pelo Teorema 12, Capítulo V, p possui uma vizinhança aberta $U = A \cap M$ (A aberto em $\mathbb{R}^{m+r}$) que é o gráfico de uma aplicação $f\colon U_0 \to \mathbb{R}^r$, de classe C^k num aberto $U_0 \subset \mathbb{R}^m$, isto é, $U = \{(x, f(x)) \in \mathbb{R}^{m+r}; x \in U_0\}$. A aplicação $\varphi\colon U_0 \to U$, $\varphi(x) = (x, f(x))$, é uma parametrização, com $p = \varphi(x_0)$, e $\Phi\colon U_0 \times \mathbb{R}^r \to \mathbb{R}^{m+r}$, $\Phi(x,y) = (x, f(x) + y)$, é um difeomorfismo de $U_0 \times \mathbb{R}^r$ sobre si mesmo, com $\Phi(U_0 \times 0) = U$. Tomamos abertos $V_0 \subset U_0$ em $\mathbb{R}^m$ e W em $\mathbb{R}^r$, tais que $x_0 \in V_0$, $0 \in W$ e $\Phi(V_0 \times W) \subset A$. Então Φ é um difeomorfismo de $V_0 \times W$ sobre o aberto $Z = \Phi(V_0 \times W) \subset \mathbb{R}^{m+r}$. Afirmamos que $\Phi(V_0 \times 0) = Z \cap M$. De fato,

$$z = \Phi(x,y) = (x, f(x) + y) \in Z \cap M \Rightarrow z \in A \cap M = \\ = U \Rightarrow y = 0 \Rightarrow z \in \Phi(V_0 \times 0).$$

Portanto, $Z \cap M \subset \Phi(V_0 \times 0)$. A inclusão oposta é evidente. Escrevamos $V = Z \cap M$ e indiquemos com $\pi\colon V_0 \times W \to V_0$ a projeção sobre o primeiro fator. A aplicação $\rho\colon Z \to V$, dada por $\rho = \varphi \circ \pi \circ \Phi^{-1}$, é uma retração de classe C^k de Z sobre V.

Demonstração do Teorema 8. Para cada $p \in M$, tomemos um aberto Z_p, uma retração $\rho_p\colon Z_p \to V_p = Z_p \cap M$, como no lema, e ponhamos $F_p = f \circ \rho_p\colon Z_p \to Y$. Para todo $x \in V_p$, temos $F_p(x) = f(\rho_p(x)) = f(x)$, logo F_p é uma extensão local de f.

Teorema 9. *Para todo $k \geq 0$, se uma aplicação $f\colon X \to \mathbb{R}^n$, definida num subconjunto arbitrário $X \subset \mathbb{R}^m$, estende-se localmente em classe C^k então f estende-se (globalmente) a uma aplicação $F\colon Z \to \mathbb{R}^n$, de classe C^k num aberto Z de $\mathbb{R}^m$ que contém X.*

Demonstração: Para cada ponto $p \in X$ temos um aberto $Z_p \subset \mathbb{R}^m$ contendo p e uma aplicação $F_p\colon Z_p \to \mathbb{R}^n$, de classe C^k, tal que $F_p(x) = f(x)$ para $x \in Z_p \cap X$. Seja $Z = \bigcup_{p \in X} Z_p$. O aberto $Z \subset \mathbb{R}^m$ é uma superfície de classe C^∞, na qual os conjuntos Z_p constituem uma cobertura

aberta. Seja $\sum\limits_{p \in X} \xi_p = 1$ uma partição da unidade estritamente subordinada a essa cobertura. Definimos $F\colon Z \to \mathbb{R}^n$ pondo $F = \Sigma\, \xi_p F_p$, ou seja $F(z) = \sum\limits_{p \in X} \xi_p(z) \cdot F_p(z)$. Evidentemente $F \in C^k$, pois $\xi_p \in C'$ e $F_p \in C^k$ para todo $p \in X$. Além disso, se $x \in X$ então

$$F(x) = \sum_p \xi_p(x) \cdot f(x), \quad = f(x),$$

logo F estende f.

Dado um subconjunto arbitrário $X \subset \mathbb{R}^m$, dizemos que uma aplicação $f\colon X \to \mathbb{R}^n$ é *diferenciável* (respect. *de classe* C^k, $1 \leq k \leq \infty$) quando f se estende localmente a uma aplicação diferenciável (respect. de classe C^k).

Observação: Excluimos o caso $k = 0$ da definição acima porque C^0 significa “contínua”e sabemos que nem toda aplicação contínua $f\colon X \to Y$ se estende localmente de modo contínuo. (Cfr. Exemplo 21.) Assim, uma aplicação $f\colon X \to \mathbb{R}^n$, de classe C^k $(k \geq 1)$, estende-se localmente, por definição, mas se $k = 0$, somente em certos casos (como, por exemplo, se X é uma superfície de classe C^1) é que existem extensões locais de classe C^0.

Sejam $X \subset \mathbb{R}^m$ e $Y \subset \mathbb{R}^n$. Uma aplicação $f\colon X \to Y$ diz-se *diferenciável* (respect. *de classe* C^k) quando, considerada como aplicação de X em $\mathbb{R}^n$ (isto é, quando composta com a inclusão $Y \to \mathbb{R}^n$), f é diferenciável (respect. de classe C^k).

Uma aplicação $f\colon X \to Y$ chama-se um *difeomorfismo* (de classe C^k) quando é diferenciável (de classe C^k) e possui uma inversa $f^{-1}\colon Y \to X$ também diferenciável (de classe C^k).

O Teorema 8 garante que a definição acima de aplicação diferenciável coincide com a que foi dada no §13 do Capítulo V no caso em que X é uma superfície no espaço euclidiano e o Teorema 9 diz que $f\colon X \to \mathbb{R}^n$ é de classe C^k $(k \geq 1)$ se, e somente se, é a restrição a X de uma aplicação de classe C^k definida numa vizinhança aberta de X em $\mathbb{R}^m$. O corolário que demonstraremos a seguir diz que o mesmo resultado vale se $\mathbb{R}^n$ for substituído pela esfera S^n, enquanto que a observação que a ele se segue informa (mas não prova) que, em vez de S^n, poder-se-ia tomar qualquer superfície de classe C^k.

Corolário do Teorema 9. *Seja* $X \subset \mathbb{R}^m$. *Se uma aplicação* $f\colon X \to$

S^n *é diferenciável* (*respect. de classe* C^k, *com* $k \geq 1$) *então* f *estende-se* (*globalmente*) *a uma aplicação diferenciável* (*respect. de classe* C^k) $F\colon Z \to S^n$ *definida num aberto* Z *de* $\mathbb{R}^m$, *contendo* X.

Com efeito, dada $f\colon X \to S^n$, podemos considerar f como uma aplicação com valores em $\mathbb{R}^{n+1}$ e, usando o Teorema 9, estendê-la a uma aplicação $G\colon W \to \mathbb{R}^{n+1}$, de classe C^k num aberto $W \subset \mathbb{R}^m$, que contém X. Seja $Z = \{z \in W; G(z) \neq 0\}$. Como $G|X = f$ toma valores em S^n, vemos que Z é um aberto de $\mathbb{R}^m$ contendo X. Consideremos a projeção radial $\rho\colon \mathbb{R}^{n+1} - \{0\} \to S^n$, $\rho(x) = x/|x|$, a qual é uma retração de classe C^∞, e definamos a aplicação $F\colon Z \to S^n$, de classe C^k, pondo $F = \rho \circ (G|Z)$, ou seja $F(z) = G(z)/|G(z)|$ par todo $z \in Z$. Vemos que F é uma extensão de f ao aberto Z de $\mathbb{R}^m$ que contém X.

Observação: O fato essencial que permitiu estender o Teorema 9 para aplicações com valores na esfera S^n é que a esfera é retrato de classe C^∞ de um aberto em $\mathbb{R}^{n+1}$. Isto é verdade em condições muito mais gerais: para toda superfície $M \subset \mathbb{R}^n$, de classe C^k, existe uma retração $\rho\colon V \to M$, de classe C^k num aberto V de $\mathbb{R}^n$, chamado uma *vizinhança tubular* de M. Daí resulta, como no corolário acima, que toda aplicação $f\colon X \to M$, classe C^k, $k \geq 1$), admite uma extensão, de classe C^k, $F\colon Z \to M$, definida num aberto Z do espaço euclidiano que contém X. A existência da vizinhança tubular é provada nos livros de Topologia Diferencial. (V. também Exerc. 13.7 do Cap. V.)

Uma advertência que deve ser feita é a seguinte: se M é uma superfície de classe C^k (digamos $k < \infty$) no espaço euclidiano, a partir da definição acima passa a ter sentido falar numa aplicação $f\colon M \to \mathbb{R}^n$ de classe C^r com $r > k$: trata-se simplesmente da restrição a M de uma aplicação de classe C^r definida num aberto do espaço euclidiano que contém M. Mas, se quisermos exprimir f em termos dos parâmetros locais de M, tais expressões não irão além da classe C^k.

Exemplo 22. Seja $f\colon [0,1] \to [0,1]$ a função identidade, $f(x) = x$. Existe, evidentemente, uma função contínua $\rho\colon \mathbb{R} \to [0,1]$ tal que $\rho|[0,1] = f$. Basta por $\rho(x) = 0$ se $x \leq 0$ e $\rho(x) = 1$ se $x \geq 1$. Noutras palavras, o intervalo $[0,1]$ é um retrato de classe C^0 da reta. Mas não existe uma função diferenciável $F\colon A \to [0,1]$, definida num aberto A, com $[0,1] \subset A \subset \mathbb{R}$, tal que $F(x) = x$ para todo $x \in [0,1]$. Com efeito, no ponto $x = 1$ a derivada de F à esquerda seria $F'(1-) = 1$ enquanto

que à direita seria

$$F'(1+) = \lim_{h\to 0^+} \frac{F(1+h)-1}{h} \le 0$$

pois $F(1+h) \in [0,1]$ para todo $h > 0$. Em geral, vê-se de modo análogo que uma bola fechada no espaço euclidiano, embora seja um retrato C^0 do espaço euclidiano, não pode ser um retrato diferenciável de um aberto que a contenha. Assim o Teorema 9, que, em virtude da Observação subseqüente, vale para aplicações $f\colon X \to M$, com valores numa superfície de classe $C^k \quad (k \ge 1)$, não vale para $f\colon X \to B$, onde B é uma bola fechada do espaço euclidiano.

Item B. *Aproximação de funções contínuas por funções diferenciáveis.*

Teorema 10. *Sejam $f\colon X \to \mathbb{R}^n$ uma aplicação contínua num conjunto $X \subset \mathbb{R}^m$ e $\varepsilon\colon X \to \mathbb{R}$ uma função contínua, com $\varepsilon(x) > 0$ para todo $x \in X$. Existe uma aplicação $g\colon A \to \mathbb{R}^n$, de classe C^∞ num aberto $A \subset \mathbb{R}^m$ contendo X, tal que $|f(x) - g(x)| < \varepsilon(x)$ para todo $x \in X$.*

Demonstração: Todo ponto $p \in X$ possui uma vizinhança $V_p = A_p \cap X$, onde $A_p \subset \mathbb{R}^m$ é aberto, tal que $|f(x) - f(p)| < \varepsilon(x)$ para todo $x \in V_p$. [Basta notar que a função contínua $\lambda_p(x) = \varepsilon(x) - |f(x) - f(p)|$ é positiva $(= \varepsilon(p))$ no ponto $x = p$, logo é positiva numa vizinhança $V_p = A_p \cap X$ desse ponto.] O aberto $A = \bigcup_{p\in X} A_p \subset \mathbb{R}^m$ é uma superfície de classe C^∞, na qual os conjuntos A_p formam uma cobertura aberta. Seja $\sum_{p\in X} \xi_p = 1$ uma partição da unidade de classe C^∞, estritamente subordinada a essa cobertura. Definamos $g\colon A \to \mathbb{R}^n$, de classe C^∞, pondo $g(x) = \sum_{p\in X} \xi_p(x) \cdot f(p)$. Para todo $x \in X$, como $f(x) = \sum_{p\in X} \xi_p(x) \cdot f(x)$, temos

$$|f(x) - g(x)| = \sum_{p\in X} \xi_p(x)|f(x) - f(p)| < \sum_{p\in X} \xi_p(x) \cdot \varepsilon(x) = \varepsilon(x).$$

A desigualdade acima se justifica porque, para todo $p \in X$, se $x \notin V_p$ então $\xi_p(x) = 0$ e, se $x \in V_p$, $\quad |f(x) - f(p)| < \varepsilon(x)$.

Corolário. *Dadas $f\colon X \to S^n$, contínua num conjunto $X \subset \mathbb{R}^m$, e $\varepsilon\colon X \to \mathbb{R}$ contínua, com $\varepsilon(x) > 0$ para todo $x \in X$, existe $g\colon A \to S^n$, de classe C^∞ num aberto $A \subset \mathbb{R}^m$ contendo X, tal que $|f(x) - g(x)| < \varepsilon(x)$ para todo $x \in X$.*

Considerando f como uma aplicação contínua de X em $\mathbb{R}^{n+1}$, o Teorema 10 fornece uma aplicação $h\colon A \to \mathbb{R}^{n+1}$, de classe C^∞ num aberto $A \subset \mathbb{R}^m$ contendo X, tal que $|f(x) - h(x)| < \varepsilon(x)/2$ para todo $x \in X$. O conjunto $A_0 = \{x \in A; h(x) \neq 0\}$ é aberto em $\mathbb{R}^m$ e contém X. Sem mudar a notação, supomos h restrita a este conjunto, logo podemos considerar a aplicação $g\colon A \to S^n$, de classe C^∞, definida por $g(x) = h(x)/|h(x)|$ para todo $x \in A$. Para todo $x \in X$, temos $|f(x) - g(x)| \leq |f(x)-h(x)|+|h(x)-g(x)|$. Concluiremos então que $|f(x)-g(x)| < \varepsilon(x)$ para todo $x \in X$, desde que mostremos que $|h(x)-g(x)| \leq |h(x)-f(x)|$. Mas isto resulta da seguinte observação elementar: $g(x) = h(x)/|h(x)|$ é o ponto da esfera S^n mais próximo do ponto $h(x)$. Como $f(x)$ pertence a S^m, segue-se que $|h(x) - g(x)| \leq |h(x) - f(x)|$.

Observações: 1) Podemos, em particular, tomar a função $\varepsilon\colon X \to \mathbb{R}$ constante. Isto equivale a aproximar a aplicação contínua $f\colon X \to \mathbb{R}^n$ por uma aplicação $g\colon X \to \mathbb{R}^n$, que é a restrição a X de uma aplicação de classe C^∞ numa vizinhança aberta de X, com $|f(g) - g(x)| < \varepsilon$ para todo $x \in X$ (aproximação uniforme). Quando X é compacto, os dois tipos de aproximação (ε uma função ou ε constante) são equivalentes, pois toda função contínua positiva $\varepsilon\colon X \to \mathbb{R}$ possui então um ínfimo positivo. No caso geral (X não compacto), entretanto, o tipo de aproximação fornecido pelo Teorema 10 é bem mais "fino"do que a aproximação uniforme, pois a função contínua positiva ε pode ter ínfimo zero em X.

2) Um resultado análogo ao do Teorema 10 (ou o seu corolário) pode ser demonstrado para aplicações $f\colon X \to N$, com valores numa superfície N de classe C^k. (Neste caso, a aproximação obtida g é apenas de classe C^k.)

3) Supondo X compacto e mantendo $\mathbb{R}^n$ como contradomínio, tem-se um resultado bem mais preciso, a saber, o Teorema de Aproximação de Weierstrass (ou sua versão mais geral, o Teorema de Stone-Weierstrass). Para uma discussão desses tópicos, veja [8].

Item C. *O Teorema do ponto fixo de Brouwer.*

Não se trata aqui de uma aplicação direta das partições da unidade, mas de uma amostra de como se utiliza o teorema de aproximação (Teorema 10) para provar resultados de Topologia.

Teorema 11 (Brouwer). *Seja B a bola fechada de centro* 0 *e raio* 1 *em $\mathbb{R}^{n+1}$. Toda aplicação contínua $f\colon B \to B$ possui (pelo menos) um*

ponto fixo.

A demonstração do Teorema 11 será feita em três etapas.

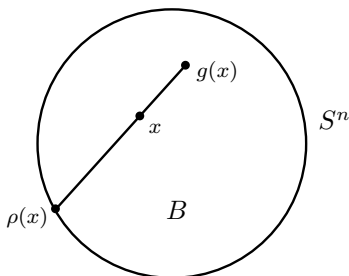
Primeira: *Se existir uma aplicação contínua $f\colon B \to B$ sem pontos fixos, existirá também uma aplicação C^∞, $g\colon B \to B$, sem pontos fixos.*

Com efeito, sendo B compacto, se for $f(x) \neq x$ para todo $x \in B$, existirá $\varepsilon > 0$ tal que $|f(x) - x| \geq \varepsilon$ para todo $x \in B$. [Basta tomar $\varepsilon =$ ínfimo da função contínua positiva $x \mapsto |f(x) - x|$ em B.] Definamos agora a aplicação contínua $h\colon B \to \mathbb{R}^{n+1}$, pondo $h(x) = (1 - \varepsilon/2)f(x)$. Temos $|h(x) - f(x)| \leq \varepsilon/2$ para cada $x \in B$. Além disso, pelo Teorema 10, existe $g\colon B \to \mathbb{R}^{n+1}$, de classe C^∞, tal que $|g(x) - h(x)| < \varepsilon/2$ para todo $x \in B$. Como $|h(x)| \leq 1 - \varepsilon/2$, vemos que $|g(x)| \leq |g(x) - h(x)| + |h(x)| \leq \varepsilon/2 + 1 - \varepsilon/2 = 1$, ou seja $g(x) \in B$ para todo $x \in B$. Para mostrar que $g\colon B \to B$ não tem pontos fixos, observemos que

$$\begin{aligned}\varepsilon \leq |x - f(x)| \leq |x - g(x)| + |g(x) - h(x)| + |h(x) - f(x)| < \\ < |x - g(x)| + \varepsilon/2 + \varepsilon/2,\end{aligned}$$

isto é, $\varepsilon < |x - g(x)| + \varepsilon$, donde $|x - g(x)| > 0$, para todo $x \in B$.

Segunda: *Se existe uma aplicação $g\colon B \to B$, de classe C^∞, sem pontos fixos, então a esfera unitária S^n é um retrato C^∞ da bola unitária $B \subset \mathbb{R}^{n+1}$.*

Com efeito, dada g, definimos a retração $\rho\colon B \to S^m$ pondo $\rho(x) =$ interseção da semi-reta $\overline{g(x)x}$ com a esfera S^n.

Evidentemente, $\rho(x) = x$ para todo $x \in S^n$. Para mostrar que ρ é de classe C^∞, notemos que $\rho(x) = g(x) + t[x - g(x)]$, onde $t \geq 1$ é único tal que $|\rho(x)| = 1$. Sendo $x \neq g(x)$ e $t \neq 0$, resulta que $\rho(x) \neq g(x)$ para todo x. Basta provarmos que t é uma função C^∞ de x, o que faremos

observando que t é definido implicitamente pela equação $F(x,t) = 1$, onde $F(x,t) = |g(x) + t[x - g(x)]|^2$. O ponto crucial a ser verificado é que $F(x,t) = 1 \Rightarrow \dfrac{\partial F}{\partial t} \neq 0$. Ora,

$$\frac{\partial F}{\partial t}(x,t) = 2\langle x - g(x), g(x) + t[x - g(x)]\rangle = 2\langle x - g(x), \rho(x)\rangle.$$

Como $g(x) - \rho(x) = -t[x - g(x)]$, vemos que $\dfrac{\partial F}{\partial t} = 0$ implica $\langle g(x) - \rho(x), \rho(x)\rangle = 0$ e daí, pelo Teorema de Pitágoras:

$$\begin{aligned} |g(x)|^2 = |g(x) - \rho(x) + \rho(x)|^2 = |g(x) - \rho(x)|^2 + |\rho(x)|^2 = \\ = |g(x) - \rho(x)|^2 + 1 > 1, \end{aligned}$$

um absurdo, pois $g(x) \in B$. [Note que $g(x) \neq \rho(x)$ para todo $x \in B$, conforme observamos acima.] Pelo Teorema da Função Implícita, $\rho \in C^\infty$.

Observação: O leitor pode exprimir t como a única raiz positiva da equação do segundo grau $|g(x) = t[x - g(x)]|^2 = 1$, com incógnita t. Assim terá uma fórmula explícita para t em função de $g(x)$.

Na terceira etapa da demonstração, seguiremos C.A. Rogers [24], que adaptou uma idéia de J. Milnor [17], por sua vez inspirada em D. Asimov.

Terceira: *Não existe uma retração $\rho\colon B \to S^n$, de classe C^∞, da bola unitária fechada $B \subset \mathbb{R}^{n+1}$ sobre a esfera S^n.*

Suponha que ρ exista. Para todo $t \in [0,1]$ definamos $\rho_t\colon B \to \mathbb{R}^{n+1}$ pondo $\rho_t(x) = x + t[\rho(x) - x] = (1-t)x + t\rho(x)$. Então $\rho_t \in C^\infty$ e, como $\rho_t(x)$ pertence ao segmento de reta $[x, \rho(x)]$, vemos que, para todo $t \in [0,1]$, $\rho_t(B) \subset B$. Seja $c = \sup_{x \in B} |\rho'(x) - \mathrm{Id}\,|$. Para $0 \leq t < 1/c$, a aplicação diferenciável $x \mapsto t[\rho(x) - x]$ tem, em todos os pontos $x \in B$, derivada com norma menor do que ou igual à constante $tc < 1$. Pelo Teorema do Valor Médio, tal aplicação é uma contração, e portanto $\rho_t\colon B \to B$ é, para $0 \leq t < 1/c$, uma perturbação da identidade, logo um difeomorfismo C^∞ de B sobre sua imagem $\rho_t(B)$. Provemos que $\rho_t(B) = B$. Seja $B = U \cup S^n$, onde U é a bola unitária aberta. Como $\rho_t(x) = x$ para todo $x \in S^n$, temos $\rho_t(S^n) = S^n$ e $\rho_t(U) \subset U$. Devemos mostrar que esta última inclusão é, na realidade, uma igualdade. Com efeito, se o aberto $\rho_t(U)$ for também fechado em U, teremos $\rho_t(U) = U$

pois U é conexo. Caso contrário, existirá $y \in U$, aderente à imagem $\rho_t(U)$ mas não pertencente a ela. Como $\rho_t(B)$ é compacto, portanto fechado, devemos ter $y \in \rho_t(B)$, isto é, $y = \rho_t(x)$ para algum $x \in B$. Como $y \notin \rho_t(U)$, não pode ser $x \in U$. Também não pode ser $x \in S^n$ pois neste caso teríamos $\rho_t(x) = x \in S^n$, enquanto $y \in U$. Assim, concluimos que, para todo $t \in \left[0, \frac{1}{c}\right)$, ρ_t é um difeomorfismo C^∞ da bola B sobre si própria. [Daí resulta que ρ_t é a restrição a B de um difeomorfismo C^∞ definido num aberto contendo a bola unitária.]

A matriz jacobiana de $\rho_t(x)$ tem a forma $I + t \cdot A(x)$, onde I é a matriz identidade $(n+1) \times (n+1)$ e $A(x)$ e a matriz jacobiana da aplicação $x \mapsto \rho(x) - x$. Portanto, o determinante jacobiano de $\rho_t(x)$ é um polinômio de grau $n+1$ em t, cujos coeficientes são funções C^∞ de x: $\det[I+t\cdot A(x)] = \sum\limits_{i=0}^{n+1} a_i(x)\cdot t^i$. Segue-se que a integral do determinante jacobiano de $\rho_t(x)$ sobre a bola B,

$$\int_B \det[I + t \cdot A(x)]dx = \sum_{i=0}^{n+1} \left[\int_B a_i(x)dx\right] t^i = p(t),$$

também é um polinômio em t (agora com coeficientes constantes). Como, para todo $t \in [0, 1/c]$, ρ_t é um difeomorfismo de B sobre si mesmo, resulta do Teorema de Mudança de Variáveis em Integrais Múltiplas que $p(t) = \text{vol}(B)$, seja qual for $t \in [0, 1/c]$. Logo, $p(t)$ é constante, qualquer que seja $t \in \mathbb{R}$. Chegaremos, então, a uma contradição se mostrarmos que $p(1) = 0$. Com efeito, para $t = 1$, ρ_t reduz-se à retração $\rho\colon B \to S^n$, cuja derivada $\rho'(x)\colon \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$ tem, para cada $x \in B$, imagem contida no subespaço vetorial de dimensão n, tangente à esfera S^n no ponto $\rho(x)$. Logo o determinante jacobiano de ρ em todos os pontos $x \in B$ é igual a zero e assim sua integral $p(1)$ é nula. Concluimos então que vol. $B = 0$, um absurdo que demonstra a inexistência de uma retração C^∞ da bola na esfera, daí resultando o Teorema do Ponto Fixo de Brouwer.

Quando provarmos o Teorema de Stokes, no parágrafo 8 mais adiante, veremos outra demonstração de que a esfera não é retrato C^∞ da bola, como caso particular de uma proposição muito mais geral.

Observação: Na terceira etapa da demonstração acima, ficou provado, na realidade, que a esfera S^n não é um retrato C^1 da bola B.

6 Integrais de superfície

Seja ω uma forma contínua de grau m, com suporte compacto, definida numa superfície m-dimensional orientada M, de classe C^1.

Como vimos no parágrafo 2, se o suporte de ω está contido na imagem de uma parametrização positiva $\varphi\colon U_0 \to U$, com $\omega(x) = a(u)du_1 \wedge \cdots \wedge du_m$ para todo $x = \varphi(u) \in U$, então a igualdade

$$\int_M \omega = \int_K a(u)du,$$

onde K é qualquer compacto J-mensurável tal que $\varphi^{-1}(\text{supp.}\,\omega) \subset K \subset U_0$, define sem ambigüidade a integral $\int_M \omega$.

Considerando os abertos $U \subset M$ e $U_0 \subset \mathbb{R}^m$ como superfícies, nas quais estão definidas as formas $\omega = adu_1 \wedge \cdots \wedge du_m$ e $\varphi^*\omega = adx_1 \wedge \cdots \wedge dx_m$, teremos ainda

$$\int_M \omega = \int_U \omega = \int_{U_0} \varphi^*\omega.$$

Observação: Se usarmos uma parametrização *negativa* $\psi\colon V_0 \to U$, relativamente à qual se tem $\omega(x) = b(v)dv_1 \wedge \cdots \wedge dv_m$, $x = \psi(v)$, então, para todo compacto J-mensurável L, tal que $\psi^{-1}(\text{supp.}\,\omega) \subset L \subset V_0$, obteremos

$$\int_L b(v)dv = -\int_M \omega.$$

Com efeito, compondo ψ com o difeomorfismo $h\colon \mathbb{R}^m \to \mathbb{R}^m$, que muda o sinal da primeira coordenada $[h(u_1, \ldots, u_m) = (-u_1, u_2, \ldots, u_m)]$, temos uma parametrização positiva $\varphi = \psi \circ h\colon U_0 \to U$, onde $U_0 = h^{-1}(V_0)$. Em relação à parametrização φ, a forma ω se escreve como $\omega(x) = -b(h(u))du_1 \wedge \cdots \wedge du_m$, $x = \varphi(u) \in U$. Como o determinante jacobiano de h tem valor absoluto igual a 1 em todos os pontos, resulta do Teorema de Mudança de Variáveis que, pondo $K = h^{-1}(L)$, vale

$$\int_L b(v)dv = \int_K b(h(u))du = -\int_K -b(h(u))du = -\int_M \omega.$$

Se as formas $\omega_1, \ldots, \omega_r$ têm suportes compactos, contidos na mesma vizinhança parametrizada U, resulta imediatamente da aditividade da integal no espaço euclidiano que se $\omega = \omega_1 + \cdots + \omega_r$ então

$$\int_M \omega = \int_M \omega_1 + \cdots + \int_M \omega_r\,.$$

A fim de definir $\int_M \omega$ no caso em que $\operatorname{supp}\omega$ é um compacto qualquer em M, tomemos uma partição da unidade $\Sigma\,\xi_i = 1$, de classe C^1, estritamente subordinada a uma cobertura aberta $M = \cup\, U_i$, onde cada U_i é imagem de uma parametrização positiva $\varphi\colon U_{i0} \to U_i$.

Para cada índice i, ponhamos $\omega_i = \xi_i \cdot \omega$. Então $\Sigma\,\omega_i = (\Sigma\,\xi_i)\cdot\omega = \omega$. O suporte de ω, sendo compacto, tem interseção não-vazia com apenas um número finito de suportes das ξ_i, pois estes constituem uma família localmente finita. Assim, $\omega_i = \xi_i \cdot \omega = 0$, salvo para um número finito de índices i, ou seja $\omega = \Sigma\,\omega_i$ é, na realidade, uma soma finita.

O suporte de cada forma ω_i está contido na vizinhança parametrizada U_i e, sendo um subconjunto fechado de $\operatorname{supp.}\omega$, é compacto. Assim, $\int_M \omega_i$ tem sentido. Pomos então, por definição,

$$\int_M \omega = \sum_i \int_M \omega_i\,.$$

Resta mostrar que a definição acima independe da partição da unidade escolhida.

Ora, se $\Sigma\,\zeta_j = 1$ é outra partição da unidade, de classe C^1, subordinada a uma cobertura aberta $M = \cup\, V_j$, onde cada V_j é imagem de uma parametrização positiva, ao lado das formas $\omega_i = \xi_i \cdot \omega$, consideremos as formas $\bar{\omega}_j = \zeta_j \cdot \omega$ e $\omega_{ij} = \xi_i \cdot \zeta_j \cdot \omega$. Temos $\sum\limits_j \omega_{ij} = \sum\limits_j \xi_i\zeta_j\omega = \xi_i(\Sigma\,\zeta_j \cdot \omega) = \xi_i \cdot \omega = \omega_i$ e, analogamente, $\sum\limits_i \omega_{ij} = \bar{\omega}_j$, para quaisquer índices i e j. Além disso, para todo i, tem-se $\operatorname{supp.}\omega_{ij} \subset U_i$ e, para todo j, tem-se $\operatorname{supp.}\omega_{ij} \subset V_j$. Logo $\int_M \omega_i = \sum\limits_j \int_M \omega_{ij}$ e $\int_M \bar{\omega}_j = \sum\limits_i \int_M \omega_{ij}$.

Daí:

$$\sum_i \int_M \omega_i = \sum_i \left[\sum_j \int_M \omega_{ij}\right] = \sum_j \left[\sum_i \int_M \omega_{ij}\right] = \sum_j \int_M \bar{\omega}_j\,,$$

o que prova ser a definição de $\int_M \omega$ independente da partição da unidade escolhida.

O teorema abaixo resume as primeiras propriedades da integral de uma forma de grau m contínua, de suporte compacto, numa superfície orientada de dimensão m.

Teorema 12. 1. $\int_M (\omega + \bar{\omega}) = \int_M \omega + \int_M \bar{\omega}$;

2. *Se* $c \in \mathbb{R}$ *então* $\int_M c \cdot \omega = c \cdot \int_M \omega$;

3. *Se* $\omega \geq 0$ *e* $\omega(x) > 0$ *para algum* $x \in M$ *então* $\int_M \omega > 0$*;*

4. *Se* ω *está definida na superfície orientada* N *e* $f\colon M \to N$ *é um difeomorfismo que preserva orientação (isto é,* M *é orientada e, para toda parametrização positiva* φ *em* M*,* $f \circ \varphi$ *é uma parametrização positiva em* N*) então* $\int_M f^*\omega = \int_N \omega$*.*

Demonstração: As igualdades 1. e 2. são evidentes quando o suporte de ω está contido numa vizinhança parametrizada. O caso geral segue-se por adição. Quanto a 3., tomando uma partição da unidade $\Sigma\,\xi_i = 1$, subordinada a uma cobertura de M por imagens de parametrizações positivas, temos $\omega = \Sigma\,\omega_i$, onde $\omega_i = \xi_i \cdot \omega \geq 0$ para todo i. Se $x \in M$ é tal que $\omega(x) > 0$, então $\omega_{i_0}(x) > 0$ para algum índice i_0. Daí, $\omega_{i_0}(y) > 0$ para todo y numa vizinhança de x, pela continuidade de ω_{i_0}. Segue-se que $\int_M \omega_i \geq 0$ para todo i e $\int_M \omega_{i_0} > 0$, pois estas são, na verdade, integrais no espaço euclidiano. Assim, $\int_M \omega = \Sigma \int_M \omega_i > 0$. Finalmente, para provar 4., escrevamos $M = \cup U_i$, onde cada U_i é imagem de uma parametrização positiva $\varphi_i\colon U_{i0} \to U_i$. Então $N = \cup f(U_i)$, onde cada $f(U_i)$ é imagem da parametrização positiva $f \circ \varphi_i$. Para calcular $\int_N \omega$, escrevemos $\omega = \Sigma\,\omega_i$, onde $\operatorname{supp.}\omega_i \subset f(U_i)$, daí resultando que $f^*\omega = \Sigma\, f^*\omega_i$, com $\operatorname{supp.} f^*\omega_i \subset U_i$ para cada i. Temos:

$$\int_N \omega_i = \int_{f(U_i)} \omega_i = \int_{U_{i0}} (f \circ \varphi_i)^*\omega_i = \int_{U_{i0}} \varphi_i^* \cdot f^*\omega_i = \\ = \int_{U_i} f^*\omega_i = \int_M f^*\omega_i$$

e, conseqüentemente,

$$\int_N \omega = \sum_i \int_N \omega_i = \sum_i \int_M f^*\omega_i = \int_M \sum_i f^*\omega_i = \int_M f^*\omega,$$

concluindo assim a demonstração do teorema.

Se $f\colon M \to N$ for um difeomorfismo que inverte a orientação (isto é, para cada parametrização positiva $\varphi\colon U_0 \to U$ em M, a parametrização $f \circ \varphi\colon U_0 \to f(U_0)$ é negativa em N) então vale a igualdade $\int_M f^*\omega = -\int_N \omega$. Isto resulta da Observação feita no início deste parágrafo.

As igualdades $\int_M f^*\omega = \int_N \omega$ para um difeomorfismo positivo $f\colon M \to N$, e $\int_M f^*\omega = -\int_N \omega$ para um difeomorfismo negativo $f\colon M \to N$, caracterizam a integral de uma forma diferencial $\int_M \omega$ sobre uma superfície orientada M como uma "integral orientada", como é a integral

curvilínea $\int_\lambda \omega$, estudada no Capítulo IV. No Apêndice a este parágrafo, veremos um caso de integral "não-orientada" numa superfície.

Se M (e conseqüentemente N) for desconexa, podemos ter um difeomorfismo $f\colon M \to N$ que preserva a orientação em algumas componentes conexas de M e inverte em outras. Neste caso, nada se pode dizer relacionando $\int_M f^*\omega$ com $\int_N \omega$. Deve-se olhar o item 4. do Teorema 12 como uma formulação global da regra de mudança de variáveis em integrais múltiplas. Formulações mais gerais, em que f não precisa ser injetiva, serão feitas no parágrafo 9.

Mais geralmente, podemos definir $\int_M \omega$ sob as seguintes hipóteses:

0. ω tem suporte compacto;

1. O conjunto dos pontos de M nos quais ω é descontínua tem medida nula;

2. ω é limitada, isto é, existe $c > 0$ tal que

$$|\omega(x) \cdot (w_1, \ldots, w_m)| \le c|w_1| \ldots |w_m|$$

para quaisquer $x \in M$ e $w_1, \ldots, w_m \in T_xM$.

Nas condições acima, para todo $x = \varphi(u)$ pertencente à imagem de uma parametrização positiva $\varphi\colon U_0 \to U$, temos $\omega(x) = a(u)du_1 \wedge \cdots \wedge du_m$, onde $a\colon U_0 \to \mathbb{R}$ é uma função cujas descontinuidades formam um conjunto de medida nula. Quando $\text{supp.}\, a = \varphi^{-1}(\text{supp.}\, \omega)$ é compacto, para que a seja integrável basta que seja limitada. Como

$$|a(u)| = \left|\omega(\varphi(u)) \cdot \left(\frac{\partial\varphi}{\partial u_1}(u), \ldots, \frac{\partial\varphi}{\partial u_m}(u)\right)\right| \le$$
$$\le c \cdot |\frac{\partial\varphi}{\partial u_1}(u)| \ldots |\frac{\partial\varphi}{\partial u_m}(u)|,$$

vemos que a será limitada se os vetores $\dfrac{\partial\varphi}{\partial u_1}(u), \ldots, \dfrac{\partial\varphi}{\partial u_m}(u)$ forem limitados em U_0, isto é, se existir uma constante c_1 tal que $|\varphi'(u)| \le c_1$ para todo $u \in U_0$.

Portanto, a integral $\int_M \omega$ pode ser definida quando ω é uma m-forma de suporte compacto em M, cumprindo as condições 1. e 2. acima. Diremos então que ω é *integrável*. Nesta situação mais geral, $\int_M \omega$ será ainda definida como $\Sigma \int_M \omega_i$, onde $\omega = \Sigma\, \omega_i$ e cada $\omega_i = \xi_i \cdot \omega$ tem suporte compacto, contido na imagem de uma parametrização positiva $\varphi_i\colon U_{i0} \to U_i$, mas agora exigiremos que a derivada de φ_i seja limitada em U_{i0}. Esta exigência é desnecessária quando ω é contínua e é

automaticamente satisfeita quando a parametrização φ_i se estende continuamente a um compacto contendo U_i .

Quando a superfície (orientável) M é compacta, toda forma contínua de grau m em M é integrável. Mais geralmente, se ω é limitada e o conjunto dos seus pontos de descontinuidade tem medida nula em M, então $\int_M \omega$ faz sentido.

Exemplo 23. Um exemplo da situação mais geral acima é o seguinte: seja ω uma forma de grau m contínua numa superfície orientada M, de dimensão m. Seja $K \subset M$ um subconjunto compacto J-mensurável. Definimos uma nova forma $\bar{\omega}$ em M, igual a ω nos pontos de K e igual a zero fora de K. Então $\bar{\omega}$ tem suporte compacto (contido em K) e seus pontos de descontinuidade (se existirem) estão contidos na fronteira de K, logo formam um conjunto de medida nula. Além disso, sendo contínua, ω é limitada no compacto K, logo $\bar{\omega}$ é limitada em M. Tem sentido portanto a integral $\int_M \bar{\omega}$ que, por definição, chamaremos a *integral de* ω *sobre o conjunto* K. Escreveremos $\int_M \bar{\omega} = \int_K \omega$. Um caso particular ocorre quando ω é o elemento de volume de M. Neste caso, $\int_K \omega$ chama-se o *volume* do conjunto K (*área* se dim. $M = 2$, *comprimento* se dim. $M = 1$). Quando M é uma hiperfície, podemos tirar proveito das formas particulares que então assume o elemento de volume ω, como vimos no fim do parágrafo 2.

Por exemplo, se K está contido na imagem de uma parametrização positiva $\varphi\colon U_0 \to U$ e $M \subset \mathbb{R}^3$ é uma superfície de dimensão 2, os pontos de U_0 são indicados por (u, v), põe-se

$$E = \left|\frac{\partial \varphi}{\partial u}\right|^2, F = \left\langle \frac{\partial \varphi}{\partial u}, \frac{\partial \varphi}{\partial v} \right\rangle, \quad G = \left|\frac{\partial \varphi}{\partial v}\right|^2 \text{ e então}$$

$$\text{área de } K = \int_{K_0} \sqrt{EG - F^2} du dv, \text{ onde } K_0 = \varphi^{-1}(K).$$

Ou (ainda no caso acima) se notarmos que as coordenadas do vetor normal unitário $v(x, y, z) = (\cos\alpha, \cos\beta, \cos\gamma)$ são os cos-senos dos ângulos que ele forma com os eixos x, y e z respectivamente, temos

$$\text{área de } K = \int_K \cos\alpha \, dy \wedge dz + \cos\beta \, dz \wedge dx + \cos\gamma \, dx \wedge dy.$$

A expressão acima é uma integral, sobre uma parte da superfície M, de uma combinação linear de $dy \wedge dz$, $dz \wedge dx$ e $dx \wedge dy$, que são formas

de grau 2 no espaço $\mathbb{R}^3$. Neste caso, os coeficientes $\cos\alpha$, $\cos\beta$ e $\cos\gamma$ só estão definidos sobre a superfície, mas isto não importa. Eles poderiam estar definidos num aberto contendo M. Sempre que tivermos uma forma $\omega = \Sigma\, a_I dx_I$, definida num aberto U do espaço euclidiano, e uma superfície $M \subset U$, de dimensão m, escreveremos $\int_M \omega$ para significar $\int_M i^*\omega$, onde $i^*\omega$ é a restrição de ω à superfície M. (Veja o Exemplo 14.)

A definição do volume de um compacto J-mensurável $K \subset M$ como a integral da forma elemento de volume de M sobre K pode ser justificada por meio das seguintes considerações geométricas. Suponhamos que $K = \varphi(B)$ seja imagem de um bloco retangular $B \subset \mathbb{R}^m$, contido no domínio de uma parametrização positiva φ. Cada partição pontilhada P^* do bloco B o decompõe em pequenos blocos B_i, em cada um dos quais foi escolhido um ponto ξ_i. A definição que demos corresponde a pôr vol. $K = \int_B f(u)du$, onde $f\colon B \to \mathbb{R}$ é definida por $f(u) = \text{vol}\left[\frac{\partial\varphi}{\partial u_1}(u), \ldots, \frac{\partial\varphi}{\partial u_m}(u)\right]$, sendo $\left[\frac{\partial\varphi}{\partial u_1}(u), \ldots, \frac{\partial\varphi}{\partial u_m}(u)\right] \subset T_xM$, $x = \varphi(u)$, o paralelepípedo que tem os vetores $\frac{\partial\varphi}{\partial u_j}(u)$ como arestas. Como sabemos, esta integral é limite de somas de Riemann $\Sigma(f;P^*) = \Sigma\, f(\xi_i) \cdot \text{vol. } B_i$. Para cada i, o bloco $B_i = [c_1e_1, \ldots, c_me_m]$ é um paralelepípedo cujas arestas são os vetores $c_1e_1, \ldots, c_me_m$. Seu volume é vol. $B_i = c_1 \ldots c_m$. O fator $f(\xi_i)$ é o volume do paralelepípedo $\varphi'(\xi_i) \cdot C$, onde $C = [e_1, \ldots, e_m]$ é o cubo unitário de $\mathbb{R}^m$. Logo a parcela $f(\xi_i) \cdot \text{vol. } B_i$ é o volume do paralelepípedo $\varphi'(\xi_i) \cdot B_i$, imagem do bloco B_i pela transformação linear $\varphi'(\xi_i)\colon \mathbb{R}^m \to T_{\varphi(\xi_i)}M$. Assim, $\Sigma(f;P^*)$ pode ser interpretada como uma soma de volumes de pequenos paralelepípedos tangentes a M, a reunião dos quais forma uma "aproximação poliedral" de K. É razoável considerar $\Sigma(f;P^*)$ como uma aproximação para vol. K, o que equivale a tomar, como definição:

$$\text{vol. } J = \lim_{|P|\to 0} \Sigma(f;P^*) = \int_K \omega.$$

No caso particular em que $m = 2$ e $K = \{(x,y,z) \in \mathbb{R}^2; (x,y) \in B, z = g(x,y)\}$ é o gráfico de uma função $g\colon B \to \mathbb{R}$, de classe C^1 no retângulo $B \subset \mathbb{R}^2$, dada a partição pontilhada P^* como acima, a soma de Riemann $\Sigma(f;P^*)$ é a soma das áreas dos paralelogramos A_i, cada um dos quais é a interseção do prisma retangular $B_i \times \mathbb{R}$, de base B_i, com o plano tangente a M no ponto $(\xi_i, g(\xi_i))$. A área de K é, portanto, o limite

da soma das áreas dos paralelogramos A_i quando a maior das bases e a maior das alturas dos retângulos B_i tendem a zero.

Esta definição geométrica da área de uma superfície (ou, mais geralmente, do volume, quando a dimensão é > 2) difere da definição do comprimento de um caminho, dada no Capítulo IV, porque lá usamos segmentos secantes e aqui usamos tangentes. A razão é a seguinte: se tentássemos definir, por exemplo, a área de uma superfície (de dimensão 2) como o limite das áreas de superfícies poliédricas nela inscritas (isto é, os vértices do poliedro são pontos pertencentes à superfície dada) veríamos que, mesmo em casos bem elementares, a área de um poliedro inscrito na superfície não possui limite quando todas as arestas das suas faces tendem a zero. Há um exemplo clássico, dado por H. Schwarz em 1881, no qual se inscreve num cilindro uma superfície poliédrica bastante simples, cuja área não converge para um limite determinado quando os lados dos triângulos que constituem suas faces tendem para zero. (Veja Exercício 6.8.) Por esse motivo, a área de uma superfície de classe C^1 é definida por meio de aproximações tangenciais. Isto, evidentemente, põe a questão de saber se é possível definir área de uma superfície de classe C^0, de modo a ter (como fizemos para caminhos) um teorema segundo o qual no caso particular de classe C^1 a área sempre existe e é dada pela fórmula acima. A teoria da área para superfícies de classe C^0 (ou mais gerais) não é de modo algum tão simples como a do comprimento de caminhos, que expusemos no Capítulo IV. Trata-se de um campo ainda com problemas em aberto. O leitor interessado pode consultar Radó [22].

As formas diferenciais são, sem dúvida, os objetos mais adequados para serem integrados sobre uma superfície. Mas, dependendo da situação, outros tipos de integrais de superfícies podem ocorrer de modo natural. Trataremos deles a seguir.

Por exemplo, que sentido tem a integral de uma função $f\colon M \to \mathbb{R}$ sobre uma superfície M (digamos, compacta e orientada)? Nos livros de Cálculo e de aplicações da Matemática, encontramos a integral $\int_M f$ definida como o limite das somas $\Sigma\, f(\xi_i) \cdot \mathrm{K}_i$ onde $\xi_i \in K_i$ para cada i, $M = \cup K_i$ é uma decomposição de M como reunião de um número finito de compactos J-mensuráveis sem pontos interiores em comum, e o limite é tomado quando a norma da decomposição (maior diâmetro dos K_i) tende para zero.

Do nosso ponto de vista, dada a função contínua $f\colon M \to \mathbb{R}$, com

suporte compacto, definida na superfície orientada M, consideramos a forma $f \cdot \omega$, produto da função f pela forma ω, elemento de volume de M, e tomamos $\int_M f = \int_M f \cdot \omega$ como definição da integral da função f sobre a superfície M. A continuidade de f é dispensável: podemos supor apenas que f é limitada, tem suporte compacto e suas descontinuidades constituem um conjunto de medida nula em M.

Exemplo 24. Seja $M \subset \mathbb{R}^{m+1}$ uma hiperfície de classe C^2, orientada por meio de um campo contínuo (e portanto C^1) de vetores normais unitários $\gamma \colon M \to S^m$. [Isto significa que uma base $\{w_1, \ldots, w_m\} \subset T_xM$ é positiva se, e somente se, $\det(\gamma(x), w_1, \ldots, w_m) > 0$, ou seja, $\{\gamma(x), w_1, \ldots, w_m\} \subset \mathbb{R}^{m+1}$ é uma base positiva.] A aplicação γ, que a cada ponto $x \in M$ associa o vetor unitário $\gamma(x)$, normal a M no ponto x, chama-se *aplicação normal de Gauss*. Para cada $x \in M$, temos $T_xM = T_{\gamma(x)}S^m$, como subespaços m-dimensionais de $\mathbb{R}^{m+1}$, ambos ortogonais ao vetor $\gamma(x)$. Logo, a derivada $\gamma'(x)$ é, na realidade, uma transformação linear $\gamma'(x) \colon T_xM \to T_xM$, de um espaço vetorial em si mesmo, e assim tem sentido o determinante $K(x) = \det.\ \gamma'(x)$, chamado a *curvatura de Gauss-Kronecker* da superfície M no ponto x. [Como $\det.\,(-\gamma'(x)) = (-1)^m \det.\ \gamma'(x)$, se tomássemos $-\gamma$ em vez de γ a curvatura gaussiana trocaria de sinal se m fosse ímpar mas independe da orientação de M quando m é par.] Sejam ω e σ, respectivamente, os elementos de volume de M e S^m. Isto implica que devemos orientar S^m, e o faremos dizendo que uma base $\{w_1, \ldots, w_m\} \subset T_pS^m$ é positiva se, e somente se, $\det(p, w_1, \ldots, w_m) > 0$, ou seja $\{p, w_1, \ldots, w_m\} \subset \mathbb{R}^{m+1}$ é uma base positiva. Desta forma, os espaços vetoriais T_xM e $T_{\gamma(x)}S^m$, não apenas são iguais como estão também munidos da mesma orientação. Então, para todo $x \in M$, temos $\omega(x) = \sigma(\gamma(x))$ pois são elementos de volume do mesmo espaço vetorial orientado. Mostraremos agora que $\gamma^*\sigma = K \cdot \omega$. Primeiro lembramos que se $A \colon E \to E$ é uma transformação linear e σ é uma forma alternada de grau igual à dimensão de E então $A^* \cdot \sigma = (\det.\ A) \cdot \sigma$. Portanto, para todo $x \in M$,

$$(\gamma^*\sigma)(x) = [\gamma'(x)]^* \cdot \sigma(\gamma(x)) = [\det.\ \gamma'(x)] \cdot \sigma(\gamma(x)) = K(x) \cdot \omega(x),$$

ou seja, $\gamma^*\sigma = K \cdot \omega$. Suponhamos que M seja compacta. A integral da função $K \colon M \to \mathbb{R}$ sobre a superfície M, ou seja, o número $\int_M K =$ (def.) $\int_M K \cdot \omega = \int_M \gamma^*\sigma$ chama-se a *curvatura integral* da superfície M. Veremos no parágrafo 9 que este número é um múltiplo inteiro do

volume da esfera S^m. Em particular, se $m = 2$, $\int_M K$ é um múltiplo inteiro de 4π.

Outro tipo de objeto que pode ser integrado sobre uma superfície M é um campo de vetores, mas agora somente em dois casos especiais.

O primeiro é quando $\dim M = 1$. Nesta situação, identificamos o vetor $F(x) = \Sigma\, a_i(x)e_i$ com a forma $\omega(x) = \Sigma\, a_i(x)dx_i$ e tomamos $\int_M F = \int_M \omega$ como definição. Isto foi feito no Capítulo IV.

Podemos também definir a integral sobre M de um campo contínuo $F\colon M \to \mathbb{R}^{m+1}$ (com suporte compacto) quando $M \subset \mathbb{R}^{m+1}$ é uma *hiperfície* orientada. Neste caso, consideraremos o campo contínuo de vetores normais unitários $v\colon M \to \mathbb{R}^{m+1}$, que define a orientação de M e pomos

$$\int_M F = \int_M \langle F, v\rangle = \int_M \langle F, v\rangle \cdot \omega,\ \omega = \text{ elemento de volume de } M.$$

Ou seja, definimos $\int_M F$ como a integral sobre M da função $x \mapsto \langle F(x), v(x)\rangle$.

Freqüentemente o campo F está definido num conjunto aberto $U \subset \mathbb{R}^{m+1}$, o qual contém a hiperfície M. A integral $\int_M F$ chama-se o *fluxo* do campo F através da hiperfície M. Se imaginarmos F como o campo das velocidades de um fluido que escoa em U sob regime estacionário (isto é, a lei de escoamento não varia com o tempo) então o fluxo $\int_M F$ mede o saldo da quantidade de fluido que atravessa M na unidade de tempo, ou seja, o que atravessa M no sentido da normal menos o que atravessa no sentido contrário.

Se $x = \varphi(u)$ pertence à imagem da parametrização positiva $\varphi\colon U_0 \to U$, sabemos que $v(x) = N(x)/|N(x)|$, onde $N(x) = \dfrac{\partial\varphi}{\partial u_1}(u) \times \cdots \times \dfrac{\partial\varphi}{\partial u_m}(u)$, e que $\omega(x) = |N(x)|du_1 \wedge \cdots \wedge du_m$. Logo

$$\langle F, v\rangle \cdot \omega = \langle F, N\rangle \cdot du_1 \wedge \cdots \wedge du_m\,.$$

Assim, se $K \subset U$ é um compacto J-mensurável, o fluxo de F através de K pode ser escrito como

$$\int_K F = \int_K \langle F, v\rangle\omega = \int_K \langle F, N\rangle du_1 \wedge \cdots \wedge du_m\,.$$

Para efeito de cálculo em termos da parametrização φ a última expressão é vantajosa, inclusive porque $\langle F, N\rangle = \text{det.}\left(F, \frac{\partial \varphi}{\partial u_1}, \ldots, \frac{\partial \varphi}{\partial u_m}\right)$, pela definição do produto vetorial.

Se $\lambda\colon M \to N$ é uma aplicação de classe C^1 de uma superfície compacta M numa superfície N e se ω é uma forma contínua em N, cujo grau é igual à dimensão de M, então podemos definir a *integral de ω ao longo da aplicação* λ pondo

$$\int_\lambda \omega = \int_M \lambda^* \omega.$$

A aplicação λ é como se fosse um "caminho a m dimensões"e a definição acima estende a noção de integral de uma forma de grau 1 ao longo de um caminho. [Estritamente falando, um intervalo $[a, b]$ não é uma superfície, mas esta discrepância será afastada no parágrafo seguinte, em que tratamos de superfícies com bordo.]

A integral usual de uma função num intervalo da reta ou num conjunto J-mensurável de $\mathbb{R}^m$ corresponde à integral de uma forma com suporte compacto. O teorema central deste capítulo (Teorema de Stokes) só vale para formas com suporte compacto. Toda forma contínua com suporte compacto é integrável. Estes são três motivos ponderáveis entre os que justificam a escolha das formas de suporte compacto como objetos adequados para a integração numa superfície. De fato, é possível desenvolver uma porção considerável dessa teoria e chegar a um conhecimento satisfatório da mesma sem integrar outras formas que não sejam as de suporte compacto.

Existem porém dois teoremas interessantes (da irrelevância dos conjuntos de medida nula e do conjunto onde a forma é zero), que usaremos apenas uma vez, no parágrafo 9, os quais não podem sequer ser enunciados, a menos que consideremos integrais de formas com suportes não-compactos. (Isto equivale, como veremos, a estudar integrais impróprias absolutamente convergentes.)

Assim, nossa definição de integral de superfície passa à terceira etapa: na primeira, tínhamos uma forma contínua com suporte compacto, contido numa vizinhança parametrizada; na segunda etapa, esta última restrição foi omitida.

Agora, definiremos a integral de uma forma contínua de grau m, numa superfície m-dimensional orientada, quer o suporte da forma seja compacto ou não.

Evidentemente, sem a restrição do suporte compacto, algumas formas contínuas são integráveis outras não. Por exemplo, a função constante, igual a 1, não é integrável na reta, nem a função contínua $y = 1/x$ é integrável no intervalo $(0, 1)$.

Seja ω uma forma contínua de grau m na superfície orientada M, de dimensão m. Tomamos uma partição da unidade $\Sigma\, \xi_i = 1$, estritamente subordinada a uma cobertura localmente finita $M = \cup\, U_i$, *onde cada* U_i *tem fecho compacto em* M e é imagem de uma parametrização positiva $\varphi_i \colon U_{i0} \to U_i$. Para cada i, pomos $\omega_i = \xi_i \cdot \omega$. As formas ω_i são contínuas, com $\text{supp.}\, \omega_i \subset \text{supp.}\, \xi_i$ compacto, contido em U_i. Temos $\omega = \Sigma\, \omega_i$ onde cada ponto $x \in M$ tem uma vizinhança na qual esta soma tem apenas um número finito de parcelas não nulas. Para cada i, tem sentido a integral $\int_M \omega_i$.

Para definir $\int_M \omega$, suponhamos inicialmente $\omega \geq 0$, donde $\omega_i \geq 0$ para todo i. [Ou seja, $\omega_i(x) = a_i(u) du_1 \wedge \cdots \wedge du_m$, com $x = \varphi_i(u)$ e $a_i(u) \geq 0$ para todo $u \in U_{i0}$.] Diremos que ω é *integrável* quando a série de números reais não-negativos $\Sigma \int_M \omega_i$ for convergente. Supondo que isto ocorre, mostraremos que a série análoga, formada com auxílio de outra partição da unidade, também converge e tem a mesma soma.

Com efeito, tomemos uma nova partição da unidade $\Sigma\, \zeta_j = 1$, subordinada à cobertura aberta localmente finita $M = \cup\, V_j$, por imagens de parametrizações positivas $\psi_j \colon V_{j0} \to V_j$, com cada $\overline{V}_j$ compacto. Sejam $\bar{\omega}_j = \zeta_j \cdot \omega$ e $\omega_{ij} = \xi_i \zeta_j \omega$. Para cada i, o compacto $\text{supp.}\, \omega_i$ intersecta apenas um número finito de conjuntos da família localmente finita $(\text{supp.}\, \omega_{ij})$. Logo, para cada i, existe apenas um número finito de índices j tais que $\omega_{ij} \neq 0$, ou seja, cada soma $\omega_i = \sum_j \omega_{ij}$ é finita. Analogamente, cada soma $\bar{\omega}_j = \sum_i \omega_{ij}$ é também finita. Podemos escrever

$$\sum_i \int_M \omega_i = \sum_i \int_M \sum_j \omega_{ij} = \sum_i \sum_j \int_M \omega_{ij}\,.$$

Como essas integrais são números ≥ 0, segue-se do Teorema 8, Capítulo 10, vol. 1 (pág. 384) que a série $\sum_j \int_M \bar{\omega}_j$ converge e

$$\sum_i \sum_j \int_M \omega_{ij} = \sum_j \sum_i \int_M \omega_{ij} = \sum_j \int_M \sum_i \omega_{ij} = \sum_j \int_M \bar{\omega}_j\,.$$

Isto mostra que, dada a forma contínua não-negativa ω em M, podemos

definir sem ambigüidade a integral de ω pela igualdade

$$\int_M \omega = \sum_i \int_M \omega_i\,,$$

desde que a série do segundo membro convirja.

No caso geral, seja ω uma forma contínua de grau m sobre a superfície orientada m-dimensional M. Escrevemos $\omega = \omega_+ - \omega_-$, onde ω_+ é a *parte positiva* e ω_- a *parte negativa* de ω, isto é, $\omega_+(x) = \omega(x)$ e $\omega_-(x) = 0$ quando $\omega(x) \geq 0$, enquanto que $\omega_-(x) = -\omega(x)$ e $\omega_+(x) = 0$ se $\omega(x) \leq 0$. Diremos então que ω é *integrável* quando as formas contínuas não-negativas ω_+ e ω_- forem ambas integráveis. Neste caso, poremos $\int_M \omega = \int_M \omega_+ - \int_M \omega_-$, por definição.

Observação 1: Se ω é integrável, $\Sigma\,\xi_i = 1$ é uma partição da unidade como acima e $\omega_i = \xi_i \cdot \omega$ então $\sum\limits_i \int_M \omega_i$ é uma série absolutamente convergente, e sua soma é $\int_M \omega$. Com efeito, as séries $\sum\limits_i \int_M \xi_i\omega_+$ e $\sum\limits_i \int_M \xi_i\omega_-$ convergem e, para cada i, temos $|\int_M \omega_i| \leq \int_M |\omega_i| = \int_M \xi_i\omega_+ + \int_M \xi_i\omega_-$. Logo $\sum\limits_i |\int_M \omega_i|$ converge. Além disso,

$$\sum \int_M \omega_i = \sum \int_M \xi_i\omega_+ - \sum \int_M \xi_i\omega_- = \int_M \omega.$$

[Em particular a soma da série independe da ordem das parcelas, o que é natural pois $\int_M \omega$ não depende da partição da unidade escolhida e a mudança da ordem corresponde a escolher outra partição, na qual as funções são as mesmas porém numeradas de forma diferente.]

Observação 2: Se $f\colon M \to N$ é um difeomorfismo de classe C^1 então, para toda forma contínua integrável ω em N, a forma induzida $f^*\omega$ é integrável em M e $\int_M f^*\omega = \pm \int_N \omega$, conforme f preserve ou inverta orientação. A demonstração se faz de modo análogo ao caso em que ω tem suporte compacto. Novamente, observa-se que se M é desconexa e f preserva orientação em algumas componentes e inverte noutras então o resultado não vale.

Observação 3. Se $U \subset \mathbb{R}^m$ é um aberto qualquer (mesmo não mensurável à Jordan), podemos considerá-lo como uma superfície m-dimensional e agora passa a ter sentido, para toda função contínua $f\colon U \to \mathbb{R}$,

a integral $\int_U f(x)dx$, que significa $\int_U \omega$, com $\omega(x) = f(x)dx_1 \wedge \cdots \wedge dx_m$. Evidentemente, ω pode ser integrável ou não. Se for, diremos que f é *absolutamente integrável* em U. Quando a função contínua f e o conjunto aberto U são ambos limitados, então $\int_U f(x)dx$ sempre existe. Com efeito, basta supor $f \geq 0$. Se $f(x) \leq c$ para todo $x \in U$ e K é um cubo em $\mathbb{R}^m$ que contém U então, para toda partição da unidade com suportes compactos $\Sigma\, \xi_i = 1$ em U e todo $k \in \mathbb{N}$ temos

$$\sum_{i=1}^{k} \int_U \xi_i f = \sum_{i=1}^{k} \int_K \xi_i f \leq c \cdot \text{vol. } K, \text{ logo a série } \sum_{i=1}^{\infty} \int_U \xi_i f$$

é convergente. [Note, porém, que a integral $\int_U \omega$ pode existir mesmo que f seja ilimitada em U.]

Observação 4: Se o aberto $U \subset \mathbb{R}^m$ é J-mensurável e $f\colon U \to \mathbb{R}$ é contínua e limitada, então a definição de $\int_U f(x)dx$ dada agora, usando partições da unidade, coincide com a que foi dada no Capítulo VI. Basta provar isto no caso em que $f \geq 0$. Sejam $\Sigma\, \xi_i = 1$ uma partição da unidade com suportes compactos em U, $f_i = \xi_i \cdot f$ e $\omega = f \cdot dx_1 \wedge \cdots \wedge dx_m$. Por definição

$$\int_U \omega = \sum_{i=1}^{\infty} \int_U f_i$$

onde, para cada i, $\int_U f_i$ significa $\int_{K_i} f_i$, com K_i J-mensurável e supp. $f_i \subset K_i \subset U$. Ora, para cada $k \in \mathbb{N}$ temos

$$\sum_{i=1}^{k} \int_U f_i = \int_U \sum_{i=1}^{k} f_i \leq \int_U f(x)dx.$$

Logo a série $\Sigma \int_U f_i$ converge, ou seja, a forma ω é integrável e $\int_U \omega \leq \int_U f(x)dx$. Para provar a desigualdade contrária, tomemos arbitrariamente um número $\varepsilon > 0$. Existe K, reunião de um número finito de blocos fechados contidos em U, tal que $\int_U f(x)dx \leq \int_K f(x)dx + \varepsilon$. A família dos suportes das funções f_i sendo localmente finita, existe $k \in \mathbb{N}$ tal que $i \geq k \Rightarrow (\text{supp. } f_i) \cap K = \varnothing$, pois K é compacto. Assim $f_1(x) + \cdots + f_k(x) = f(x)$ para todo $x \in K$ e daí

$$\int_K f(x)dx \leq \int_U \sum_{i=1}^{k} f_i \leq \int_U \omega.$$

Segue-se que $\int_U f(x)dx \leq \int_U \omega + \varepsilon$. Como ε é arbitrário, concluimos que $\int_U f(x)dx \leq \int_U \omega$, o que prova a afirmação feita.

Observação 5: Se a forma contínua ω é integrável em M então ω é integrável em qualquer aberto $A \subset M$ (considerado como superfície). Se, além disso, $\omega \geq 0$ então $\int_A \omega \leq \int_M \omega$. Suponhamos, inicialmente, $\omega \geq 0$. Seja $i \colon A \to M$ a inclusão. Mediante uma partição da unidade, escrevamos a soma localmente finita

$$i^*\omega = \sum_{i=1}^{\infty} \omega_i \,,$$

onde cada ω_i é ≥ 0, contínua, com suporte compacto, contido em A. Para cada $k \in \mathbb{N}$ podemos considerar $\alpha_k = \omega_1 + \cdots + \omega_k$ como uma forma contínua de suporte compacto em M, pondo-a igual a zero em $M - A$. Portanto, α_k é integrável em M, com

$$\int_M \alpha_k \leq \int_M \omega. \text{ Logo } \sum_{i=1}^{k} \int_A \omega_i = \int_A \alpha_k = \int_M \alpha_k \leq \int_M \omega$$

para todo $k \in \mathbb{N}$. Assim, a série $\Sigma \int_A \omega_i$ converge para o valor $\int_A \omega \leq \int_M \omega$. Se ω não é ≥ 0, pomos $\omega = \omega_+ - \omega_-$ onde cada parcela é ≥ 0, logo integrável em A, portanto ω é integrável em A. [Mas agora não podemos concluir que $\int_A \omega \leq \int_M \omega$.]

Na demonstração do teorema seguinte, vamos utilizar o

Lema 1. *Sejam ω contínua integrável em M e $F \subset M$ um subconjunto fechado de medida nula. Para todo $\varepsilon > 0$ existe Z aberto em M, contendo F, tal que $|\int_Z \omega| < \varepsilon$.*

Demonstração: Basta considerar o caso em que $\omega \geq 0$. Pelo Lema 3, usado na demonstração do Teorema 7 (§4), podemos cobrir M e (conseqüentemente) F com uma família localmente finita de abertos $U_i = \varphi_i(B(3))$ tais que os correspondentes $W_i = \varphi_i(B(1))$ ainda cobrem F. Tomamos apenas os U_i tais que $W_i \cap F \neq \varnothing$. Para cada i, seja $K_i = \overline{W}_i \cap F$. Os K_i são compactos de medida nula, com $\cup K_i = F$. Os abertos $V_i = \varphi_i(B(2))$ formam uma cobertura localmente finita de F, com cada $\overline{V}_i$ compacto, $K_i \subset V_i$ e $L_i = \varphi_i^{-1}(K_i) \subset B(2)$ é um compacto de medida nula, para cada i. Seja $\omega(x) = a_i(v)dv_1 \wedge \cdots \wedge dv_m$ a expressão de ω nos pontos $x = \varphi_i(v) \in V_i$, dada pela parametrização positiva φ_i. Cada uma das funções contínuas $a_i \colon B(2) \to \mathbb{R}$ é limitada, pois se estende

continuamente ao compacto $\overline{B(2)} \subset B(3)$. Seja $a_i(v) \leq c_i$ para todo $v \in B(2)$. Podemos cobrir L_i com um número finito de cubos abertos cuja soma dos volumes é menor do que $\varepsilon / c_i \cdot 2^i$. Seja Z_{i0} a reunião desses cubos. Ponhamos $Z_i = \varphi_i(Z_{i0})$ e $Z = \cup Z_i$. A cobertura $Z = \cup Z_i$ é localmente finita e cada Z_i tem fecho compacto. Tomando uma partição da unidade $\Sigma \xi_i = 1$ em Z, com supp. $\xi_i \subset Z_i$, vemos que

$$\omega_i = \xi_i \cdot \omega = \xi_i \cdot a_i \cdot dv_1 \wedge \cdots \wedge dv_m \leq a_i \cdot dv_1 \wedge \cdots \wedge dv_m ,$$

logo $\int_Z \omega_i \leq \int_{Z_{i0}} a_i(v) dv_1 dv_2 \ldots dv_m \leq c_i \cdot \mathrm{vol}(Z_{i0}) < \varepsilon / 2^i$.

Portanto $\displaystyle\int_Z \omega = \sum_i \int_Z \omega_i < \sum_{i=1}^{\infty} \frac{\varepsilon}{2^i} = \varepsilon.$

Teorema (Irrelevância dos conjuntos de medida nula.) *Se ω é uma forma contínua de grau m, integrável na superfície m-dimensional orientada M, e $F \subset M$ é um subconjunto fechado de medida nula, então*

$$\int_M \omega = \int_{M-F} \omega.$$

Demonstração: Sem perda de generalidade, podemos supor que $\omega \geq 0$. Dado arbitrariamente $\varepsilon > 0$, tomemos em M um aberto $Z \supset F$, tal que $\int_Z \omega < \varepsilon$, conforme o Lema 1. Seja $A = M - F$. Tomando uma partição da unidade $\xi_A + \xi_Z = 1$ em M, estritamente subordinada à cobertura $M = A \cup Z$, podemos escrever

$$\begin{aligned} \int_A \omega \leq \int_M \omega &= \int_M \xi_A \cdot \omega + \int_M \xi_Z \cdot \omega = \\ = \int_A \xi_A \cdot \omega + \int_Z \xi_Z \cdot \omega &\leq \int_A \omega + \int_Z \omega < \int_A \omega + \varepsilon. \end{aligned}$$

Isto sendo válido para todo $\varepsilon > 0$, concluimos que $\int_M \omega = \int_A \omega$.

Corolário. *Sejam ω uma forma contínua integrável na superfície orientada M e $\varphi \colon U_0 \to U$ uma parametrização positiva em M, tal que $M - U$ tem medida nula. Então $\int_M \omega = \int_U \varphi^* \omega$.*

O corolário acima é útil para o cálculo de integrais em superfícies. A esfera, o toro e todas as superfícies compactas (orientadas) em $\mathbb{R}^3$ admitem uma parametrização do tipo acima. Por exemplo, retirando-se um meridiano e um paralelo de um toro de revolução, o aberto restante

pode ser facilmente parametrizado. [Isto é verdade também para *qualquer* superfície compacta orientada mas a demonstração em dimensão > 2 requer técnicas bastante avançadas de Geometria Diferencial.]

Na integral de uma forma sobre uma superfície também podemos desprezar qualquer conjunto fechado no qual a forma se anule. Isto é o

Teorema (Irrelevância dos zeros.) *Sejam ω uma forma contínua, integrável na superfície orientada M, e $F \subset M$ um subconjunto fechado tal que $\omega(x) = 0$ para todo $x \in F$. Então $\int_M \omega = \int_{M-F} \omega$.*

A demonstração se faz tal qual a do teorema anterior, utilizando-se, em vez do Lema 1, o

Lema 2. *Seja $F \subset M$ um conjunto fechado, em todos os pontos do qual a forma contínua ω (de grau m na superfície m-dimensional orientada M) se anula. Dado qualquer $\varepsilon > 0$, existe um aberto Z em M, contendo F, tal que $|\int_Z \omega| < \varepsilon$.*

Demonstração: Basta supor $\omega \geq 0$. Tomamos uma família localmente finita de abertos $U_i = \varphi_i(B(3))$ tais que os $W_i = \varphi_i(B(1))$ cobrem F e $W_i \cap F \neq \varnothing$ para todo i. [Como no Lema 1.] Pomos $K_i = \overline{W}_i \cap F$. Os K_i são compactos, com $\cup K_i = F$, os abertos $V_i = \varphi_i(B(2))$ formam uma cobertura localmente finita de F, cada um deles tem fecho compacto em M, é imagem da parametrização positiva $\varphi_i \colon B(2) \to V_i$ e contém K_i. Seja $\omega(x) = a_i(v) dv_1 \wedge \cdots \wedge dv_m$ a expressão de ω nos pontos $x = \varphi_i(v) \in V_i$, dada pela parametrização φ_i. Temos $a_i(v) = 0$ para cada $v \in L_i = \varphi_i^{-1}(K_i) \subset B(2)$. Seja c_i o volume exterior do compacto L_i. (Note que $c_i = \inf. \Sigma \operatorname{vol}(C_j)$, onde $L_i \subset C_1 \cup \cdots \cup C_k$ é uma cobertura de L_i por cubos abertos.) Tomemos um aberto $Z_{i0} \subset B(2)$, reunião finita de cubos abertos contendo L_i, cuja soma dos volumes é $< 2c_i$, tal que $a_i(v) < \varepsilon/2^{i+1} c_i$ para todo $v \in Z_{i0}$. Então $\operatorname{vol}(Z_{i0}) < 2c_i$, logo $\int_{Z_{i0}} a_i(v) dv < \varepsilon/2^i$. Sejam $Z_i = \varphi_i(Z_{i0})$ e $Z = \cup Z_i$. Seja $\Sigma \xi_i = 1$ uma partição da unidade em Z, estritamente subordinada à cobertura $Z = \cup Z_i$. Temos

$$\int_Z \omega = \sum_i \int_{Z_{i0}} \xi_i(v) a_i(v) dv < \sum_i \frac{\varepsilon}{2^i} = \varepsilon,$$

como queríamos demonstrar.

As considerações acima desenvolvidas podem ser aplicadas para estender o Teorema de Mudança de Variáveis, eliminando a hipótese de

que o determinante jacobiano da mudança de variáveis seja sempre diferente de zero, conforme o

Exemplo 25. Sejam $U, V \subset \mathbb{R}^m$ abertos e $h\colon U \to V$ um homeomorfismo de classe C^1. Seja $f\colon V \to \mathbb{R}$ uma função contínua limitada. Então $\int_K f(h(x))|J(x)|dx = \int_\Omega f(y)dy$, onde $J(x) = \det.\ h'(x)$. Com efeito, $S = \{x \in U; J(x) = 0\}$ é um subconjunto fechado de U, cuja imagem $h(S)$ é um subconjunto fechado de V, com medida nula, pelo Teorema de Sard. Consideremos a m-forma ω em V, definida por $\omega(y) = f(y)dy_1 \wedge \cdots \wedge dy_m$. Temos $(h^*\omega)(x) = f(h(x)) \cdot J(x)dx_1 \wedge \cdots \wedge dx_m$ para todo $x \in U$. Como $h\colon U - S \to V - h(S)$ é um difeomorfismo, se supusermos $J(x) \geq 0$ para todo $x \in U$ teremos

$$\int_U f(h(x)) \cdot |J(x)|dx = \int_U h^*\omega = \int_{U-S} h^*\omega = \int_{V-h(S)} \omega = \\ = \int_V \omega = \int_V f(y)dy.$$

Se admitirmos porém que seja $J(x) \leq 0$ para todo $x \in U$ então $h\colon U - S \to V - h(S)$ é um difeomorfismo que inverte orientação, logo $\int_{U-S} h^*\omega = -\int_{V-h(S)} \omega$. Além disso, neste caso será $|J(x)| = -J(x)$. Fazendo estas duas mudanças de sinal na seqüência de igualdades acima, obteremos ainda no final

$$\int_U f(h(x)) \cdot |J(x)|dx = \int_V f(y)dy.$$

Em geral, podemos ter $J(x) > 0$ em alguns pontos de U e, noutros pontos $J(x) < 0$. Então escrevemos $U = A \cup B$, reunião disjunta de dois abertos, onde $A = \{x \in U; J(x) > 0\}$ e $B = U - A$. Segue-se que $V = h(A) \cup h(B)$, também reunião disjunta de abertos e, pelo que vimos acima, temos

$$\int_U f(h(x)) \cdot |J(x)|dx = \int_A f(h(x))|J(x)|dx + \\ + \int_B f(h(x)) \cdot |J(x)|dx = \int_{h(A)} f(y)dy + \int_{h(B)} f(y)dy = \int_V f(y)dy.$$

APÊNDICE

Examinemos agora a hipótese de orientabilidade da superfície sobre a qual se efetua uma integração. Ela é, de fato, necessária para se definir a integral de uma forma diferencial e constitui um dado importante nas aplicações desta noção. (Veremos alguns exemplos disso nos parágrafos finais deste capítulo.) Somente para citar um caso simples, mencionaremos que o fluxo de um campo de vetores através de uma hiperfície só tem sentido quando a mesma é orientável.

Por outro lado, é intuitivamente óbvio que uma faixa de Möbius tem área (igual à área do retângulo que foi torcido e colado para formá-la). Este exemplo indica que a orientabilidade é uma exigência que deveria ser evitada na definição geral de volume de uma superfície. Ainda sobre a faixa de Möbius M: se a densidade de M em cada um dos seus pontos x tem um valor $f(x)$, então a massa de M deveria ser expressa como uma integral $\int_M f$, embora M não seja orientável.

As considerações acima sugerem um reexame da definição de integral, visando retirar a condição de orientabilidade. Isto nos conduz à noção de densidade escalar, como objeto que pode ser integrado, mesmo em superfície não-orientáveis.

Uma *densidade escalar* num espaço vetorial m-dimensional E é uma função $\rho\colon E \times \cdots \times E$ (m fatores) $\to \mathbb{R}$ com a seguinte propriedade: se $A\colon E \to E$ é uma transformação linear e $v_1, \ldots, v_m \in E$ então

$$\rho(A{\cdot}v_1, \ldots, A{\cdot}v_m) = |\det.\, A| \cdot \rho(v_1, \ldots, v_m).$$

Por exemplo, se ω é qualquer m-forma alternada em E então $|\omega|$ é uma densidade escalar. Isto mostra que em todo espaço vetorial $E \neq \{0\}$ existem densidades escalares não-nulas.

Um importante exemplo de densidade escalar é a *densidade de volume* ρ, num espaço vetorial E, munido de um produto interno. (Não supomos E orientado!) Por definição, dados $v_1, \ldots, v_m \in E$, tem-se

$$\rho(v_1, \ldots, v_m) = \sqrt{\det.\,(\langle v_i, v_j\rangle)} = \text{vol.}\,[v_1, \ldots, v_m],$$

onde $[v_1, \ldots, v_m] \subset E$ é o paralelepípedo que tem por arestas os vetores v_i. Se dotarmos E de uma orientação e chamarmos de ω a forma elemento de volume correspondente, teremos $\rho = |\omega|$.

O conjunto $\Delta(E)$ das densidades escalares no espaço vetorial E, com as operações usuais de soma de funções e produto por um número real, é um espaço vetorial.

Seja $\{e_1, \ldots, e_m\}$ uma base no espaço vetorial E. Uma vez fixada esta base, indicaremos com a notação $\det.(v_1, \ldots, v_m)$, para $v_1, \ldots, v_m \in E$, o determinante da matriz cuja j-ésima coluna é formada pelas coordenadas do vetor v_j em relação à base dada. Ou ainda, $\det(v_1, \ldots, v_m)$ é o determinante da transformação linear $A\colon E \to E$ tal que $A \cdot e_1 = v_1, \ldots, A \cdot e_m = v_m$.

Isso posto, para toda densidade escalar $\rho \in \Delta(E)$ existe um número real a tal que $\rho(v_1, \ldots, v_m) = |\det(v_1, \ldots, v_m)| \cdot a$. Basta pôr $a = \rho(e_1, \ldots, e_m)$. Desta observação seguem-se os seguintes fatos, relativos a uma densidade escalar ρ no espaço vetorial E, de dimensão m:

1. Se $v_1, \ldots, v_m$ são linearmente dependentes então $\rho(v_1, \ldots, v_m) = 0$. Em particular, se a lista $v_1, \ldots, v_m$ tem repetições, $\rho(v_1, \ldots, v_m) = 0$;
2. Se $\rho \neq 0$ e $v_1, \ldots, v_m$ são linearmente independentes então $\rho(v_1, \ldots, v_m) \neq 0$;
3. $\rho(v_1, \ldots, c \cdot v_i, \ldots, v_m) = |c| \rho(v_1, \ldots, v_i, \ldots, v_m)$;
4. Se $i \neq j$ então $\rho(v_1, \ldots, v_i + c \cdot v_j, \ldots, v_m) = \rho(v_1, \ldots, v_i, \ldots, v_m)$, isto é, o valor $\rho(v_1, \ldots, v_m)$ não se altera se somarmos a uma de suas variáveis um múltiplo qualquer de outra;
5. $\rho(v_1, \ldots, v_m)$ é uma função simétrica, isto é, se σ é qualquer permutação dos inteiros $1, \ldots, m$ então
$$\rho(v_{\sigma(1)}, \ldots, v_{\sigma(m)}) = \rho(v_1, \ldots, v_m);$$
6. O espaço vetorial $\Delta(E)$ tem dimensão 1.

Com efeito, seja $\rho_0 \neq 0$. Dada qualquer $\rho \in \Delta(E)$, tomemos uma base $\{e_1, \ldots, e_m\}$ em E e seja $c = \rho(e_1, \ldots, e_m)/\rho_0(e_1, \ldots, e_m)$. Então $\rho = c \cdot \rho_0$, logo $\{\rho_0\}$ é uma base de $\Delta(E)$.

Diremos que uma densidade escalar ρ é *positiva* quando existirem $e_1, \ldots, e_m \in E$ tais que $\rho(e_1, \ldots, e_m) > 0$. Então $\rho(v_1, \ldots, v_m) > 0$ para toda lista de vetores linearmente independentes $v_1, \ldots, v_m$.

Sejam E, F espaços vetoriais de mesma dimensão m. Toda transformação linear $T\colon E \to F$ induz uma transformação linear $T^*\colon \Delta(F) \to \Delta(E)$, que associa à densidade escalar ρ em F a densidade escalar $T^*\rho$ em E, definida por:

$$(T^*\rho)(u_1, \ldots, u_m) = \rho(T \cdot u_1, \ldots, T \cdot u_m).$$

Resta mostrar que, de fato $T^*\rho \in \Delta(E)$. Isto é claro se T não for um isomorfismo pois, neste caso, $T \cdot u_1, \ldots, T \cdot u_m$ são sempre linearmente dependentes, logo $T^*\rho = 0$. Se, porém, $T^{-1}\colon F \to E$ existe, escrevamos $u_1 = T^{-1} \cdot v_1, \ldots, u_m = T^{-1} \cdot v_m$. Então, dada uma transformação linear $A\colon E \to E$, temos

$$
\begin{aligned}
(T^*\rho) \cdot (A \cdot u_1, \ldots, A \cdot u_m) &= (T^*\rho)(AT^{-1} \cdot v_1, \ldots, AT^{-1} \cdot v_m) = \\
&= \rho(TAT^{-1} \cdot v_1, \ldots, TAT^{-1} \cdot v_m) = \\
&= |\det.\ TAT^{-1}| \cdot \rho(v_1, \ldots, v_m) = \\
&= |\det.\ A| \cdot \rho(T \cdot u_1, \ldots, T \cdot u_m) = \\
&= |\det.\ A| \cdot (T^*\rho)(u_1, \ldots, u_m).
\end{aligned}
$$

Evidentemente, $T^*(\rho_1 + \rho_2) = T^*\rho_1 + T^*\rho_2$ e $T^*(c\rho) = c \cdot T^*\rho$, logo $T^*\colon \Delta(F) \to \Delta(E)$ é uma transformação linear, que é igual a zero se T não for bijetiva e é isomorfismo se T o for. Além disso, neste último caso, $\rho > 0$ em $\Delta(E)$ implica $T^*\rho > 0$. Observemos ainda que se ST faz sentido então $(ST)^* = T^*S^*$.

A densidade escalar $T^*\rho$ diz-se *induzida* por ρ em E, por meio de T. Diz-se também que $T^*\rho$ é o *pull-back* da densidade ρ pela transformação T.

No Cálculo Tensorial clássico, uma densidade escalar é uma função ρ que associa a cada base $\{e_1, \ldots, e_m\}$ um número real a, chamado a coordenada de ρ na base dada [no caso, $a = \rho(e_1, \ldots, e_m)$] de tal modo que se $\{f_1, \ldots, f_m\}$ é outra base, à qual ρ faz corresponder o número b então

$$b = |\det.\ A| \cdot a,$$

onde $A\colon E \to E$ é a transformação linear que leva a base $\{e_1, \ldots, e_m\}$ na base $\{f_1, \ldots, f_m\}$. Se $f_j = \sum_{i=1}^{m} \alpha_{ij} e_i$ $(j = 1, \ldots, m)$ então det. A é o determinante da matriz (α_{ij}). Evidentemente, temos também $a = |\det.\ B| \cdot b$, onde $B \cdot f_j = e_j$, com $e_j = \Sigma\, \beta_{ij} f_i$ e det. $B = \det.(\beta_{ij})$.

Uma *densidade escalar numa superfície* M é uma aplicação ρ que a cada ponto $x \in M$ associa uma densidade $\rho(x) \in \Delta(T_x M)$.

Dada uma densidade escalar ρ em M, a cada parametrização $\varphi\colon U_0 \to U$ em M corresponde uma função real $a\colon U_0 \to \mathbb{R}$, definida por

$$a(u) = \rho(x) \cdot \left(\frac{\partial \varphi}{\partial u_1}(u), \ldots, \frac{\partial \varphi}{\partial u_m}(u) \right), \quad x = \varphi(u).$$

Se $\psi\colon V_0 \to V$ é outra parametrização, com $U \cap V \neq \varnothing$, e a ela corresponde a função $b\colon V_0 \to \mathbb{R}$, então, para cada $x = \varphi(u) = \psi(v) \in U \cap V$, vale

$$a(u) = |J(u)| \cdot b(v),$$

onde $J(u)$ é o determinante jacobiano do difeomorfismo $\psi^{-1} \circ \varphi\colon \varphi^{-1}(U \cap V) \to \psi^{-1}(U \cap V)$, calculado no ponto u.

A igualdade acima pode servir de base a uma definição, segundo a qual uma densidade escalar numa superfície M é uma aplicação que a cada parametrização $\varphi\colon U_0 \to U$ em M associa uma função real $a\colon U_0 \to \mathbb{R}$, de tal modo que se $b\colon V_0 \to \mathbb{R}$ é a função associada à parametrização $\psi\colon V_0 \to V$, com $U \cap V \neq \varnothing$, então, para todo $x = \varphi(u) = \psi(v) \in U \cap V$ tem-se $a(u) = |J(u)| \cdot b(v)$, onde $J(u) = \det.[(\psi^{-1} \circ \varphi)'(u)]$. Esta definição equivale à que demos. Para verificar isto, reobtenha a densidade $\rho(x)$ a partir da função a pondo

$$\rho(x)\cdot\left(\frac{\partial\varphi}{\partial u_1}(u), \ldots, \frac{\partial\varphi}{\partial u_m}(u)\right) = a(u)$$

se $x = \varphi(u)$ e $\rho(x) \cdot (v_1, \ldots, v_m) = |\det. A| \cdot a(u)$, onde $A = (\alpha_{ij})$ e

$$v_j = \sum_i \alpha_{ij} \frac{\partial\varphi}{\partial u_i}(u) \in T_xM.$$

A relação $a(u) = |J(u)| \cdot b(v)$ implica que a definição de $\rho(x)$ não depende da parametrização φ e que $\rho(x)$ é, de fato, uma densidade escalar em T_xM.

Se M é de classe C^k, tem sentido dizer que ρ é uma *densidade escalar de classe* C^s em M, para $0 \leq s < k$. Isto significa que cada ponto $x \in M$ pertence à imagem de uma parametrização $\varphi\colon U_0 \to U$, de classe C^k em M, tal que a função $a\colon U_0 \to \mathbb{R}$, definida por

$$a(u) = \rho(x)\cdot\left(\frac{\partial\varphi}{\partial u_1}(u), \ldots, \frac{\partial\varphi}{\partial u_m}(u)\right),$$

$x = \varphi(u)$, é de classe C^s. Então a regra de transformação $a(u) = |J(u)| \cdot b(v)$ mostra que, para qualquer outra parametrização $\psi\colon V_0 \to V$, de classe C^k em M, a função $b(v)$ também é de classe C^s, pois o jacobiano $J(u)$ é uma função de classe C^{k-1}.

Uma densidade escalar ρ em M diz-se *positiva* (escreve-se $\rho > 0$) quando $\rho(x) > 0$ para todo $x \in M$. Isto significa que para qualquer

lista $v_1, \dots, v_m \in T_xM$ de vetores linearmente independentes, tem-se $\rho(x) \cdot (v_1, \dots, v_m) > 0$. Note-se que M não precisa ser orientável para que esta definição tenha sentido.

Daremos agora alguns exemplos. Se ρ é uma densidade de classe C^0 em M então $|\rho|$ é também uma densidade de classe C^0. (Mas não é de classe C^r, em geral, mesmo que ρ o seja, salvo, é claro, quando $\rho(x) \neq 0$ para todo x.) Se ω é uma forma de grau máximo em M então $|\omega|$ é uma densidade escalar, de classe C^0 se $\omega \in C^0$. Seja $\mathcal{C}$ uma cobertura de M por abertos U, em cada um dos quais está definida uma forma diferencial ω_U de grau máximo. Suponha que, para cada $x \in U \cap V$ se tenha $|\omega_U(x)| = |\omega_V(x)|$. Então obteremos uma densidade escalar ρ em M pondo $\rho(x) = |\omega_U(x)|$, onde U é qualquer aberto da cobertura $\mathcal{C}$ contendo o ponto x. Se cada ω_U é de classe C^0 então $\rho \in C^0$. Se cada ω_U é de classe C^r e $\omega_U(x) \neq 0$ para todo $x \in U$ então $\rho \in C^r$.

O exemplo mais importante de densidade escalar numa superfície M, de classe C^k, é a *densidade de volume* ρ, definida por

$$\rho(x) \cdot (v_1, \dots, v_m) = \text{vol.}\,[v_1, \dots, v_m]$$

para quaisquer $x \in M$ e $v_1, \dots, v_m \in T_xM$. Acima, $\text{vol.}\,[v_1, \dots, v_m]$ é o volume (não-orientado!) do paralelepípedo cujas arestas são os vetores $v_1, \dots, v_m$. Dada uma parametrização $\varphi \colon U_0 \to U$, de classe C^k em M, pomos

$$g_{ij}(u) = \left\langle \frac{\partial \varphi}{\partial u_i}(u), \frac{\partial \varphi}{\partial u_j}(u) \right\rangle, \quad u \in U_0\,,$$

e $g(u) = (g_{ij}(u))$ (matriz de Gram dos vetores básicos $\frac{\partial \varphi_i}{\partial u_j}$ no ponto u). Então

$$\rho(x) \cdot \left(\frac{\partial \varphi}{\partial u_1}(u), \dots, \frac{\partial \varphi}{\partial u_m}(u) \right) = \sqrt{\text{det. } g(u)}, \quad x = \varphi(u),$$

logo $a(u) = \sqrt{\text{det. } g(u)}$, o que mostra que a densidade de volume é de classe C^{k-1}.

Quando a superfície M é orientada e seu elemento de volume é a forma ω então a densidade de volume de M é $|\omega|$. No caso geral, tomamos a cobertura $\mathcal{C}$ de M por imagens de parametrizações $\varphi \colon U_0 \to U$, de classe C^k e consideramos cada aberto U como uma superfície orientada, na qual $\varphi > 0$. Seja ω_U o elemento de volume correspondente. Então $|\omega_U(x)| = |\omega_V(x)|$ para todo $x \in U \cap V$ e a densidade de volume ρ é tal que $\rho(x) = |\omega_U(x)|$ seja qual for U contendo x.

Sejam M, N superfícies de classe C^k, com a mesma dimensão e $f: M \to N$ uma aplicação de classe C^k. Se ρ é uma densidade escalar de classe C^r $(r < k)$ em N então a densidade $f^*\rho$, de classe C^r em M, *induzida por* ρ *em* M, por meio de f é definida pondo-se, para cada $x \in M$ e $w_1, \ldots, w_m \in T_xM$,

$$[(f^*\rho)(x)](w_1, \ldots, w_m) = \rho(f(x)) \cdot (f'(x) \cdot w_1, \ldots, f'(x) \cdot w_m).$$

A transformação linear $f'(x): T_xM \to T_{f(x)}N$, derivada de f no ponto x, induz uma transformação $[f'(x)]^*: \Delta(T_{f(x)}N) \to \Delta(T_xM)$. Tem-se $(f^*\rho)(x) = [f'(x)]^*\rho(f(x))$, logo $f^*\rho$ é uma densidade escalar em M. Se consideramos outra aplicação $g: N \to P$, veremos facilmente que $(g \circ f)^*\rho = f^*(g^*\rho)$ para toda densidade ρ na superfície P.

Revendo (desde sua primeira parte, na seção 2) a definição da integral de uma forma contínua de grau máximo e suporte compacto numa superfície orientada M, vemos que as mesmas considerações se aplicam, quase ipsis literis, ao caso em que não se supõe M orientável mas, em lugar da forma ω, usamos uma densidade escalar ρ. O único ponto em que é preciso mudar alguma coisa é na prova de que a definição $\int_M \rho = \int_K a(u)du$ não depende da parametrização φ escolhida [no caso em que ρ tem suporte compacto, contido na imagem da parametrização $\varphi: U_0 \to U$ e $\varphi^{-1}(\mathrm{supp.}\,\omega) \subset K \subset U_0$].

Se $\psi: V_0 \to V$ é outra parametrizaçção com $\mathrm{supp.}\,\rho \subset V$ e

$$b(v) = \rho(x)\left(\frac{\partial\psi}{\partial v_1}(v), \ldots, \frac{\partial\psi}{\partial v_m}(v)\right), \quad x = \psi(v) = \varphi(u) \in U \cap V,$$

sabemos que $a(u) = |J(u)| \cdot b(v)$. Logo, pondo $L = (\psi^{-1} \circ \varphi)(K)$, temos

$$\int_L b(v)dv = \int_K b(v) \cdot |J(u)|du = \int_K a(u)du,$$

onde a primeira igualdade deve-se ao Teorema de Mudança de Variáveis. [A única alteração relativamente ao caso de formas consistiu em usar $|J(u)|$ em vez de $J(u)$, pois agora não sabemos se $J(u) > 0$ para todo $u \in U \cap V$.]

Usando as partições da unidade, como antes, definimos $\int_M \rho$ onde ρ é uma densidade contínua com suporte compacto na superfície M, que pode ser orientada ou não. Mais geralmente, $\int_M \rho$ tem sentido quando ρ tem suporte compacto, é limitada e o conjunto dos seus pontos de descontinuidade tem medida nula em M.

A integral de uma densidade escalar é "não-orientada", isto é, se $f: M \to N$ é um difeomorfismo e ρ é uma densidade contínua com suporte compacto em N, então $f^*\omega$ é uma densidade contínua com suporte compacto em M e vale $\int_M f^*\rho = \int_N \rho$. Mesmo que M e N sejam orientadas e que $f: M \to N$ seja um difeomorfismo que inverte orientações, temos ainda a igualdade $\int_M f^*\rho = \int_N \rho$, para toda densidade escalar ρ em N.

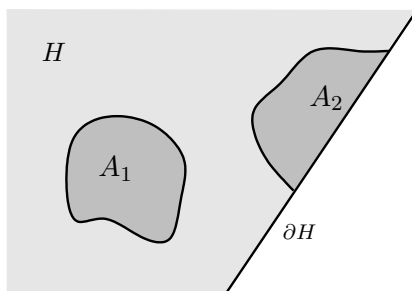
Agora tem sentido definir o volume de um subconjunto compacto J-mensurável $K \subset M$, mesmo que a superfície M não seja orientável. Tomamos a densidade de volume ρ de M, definimos uma nova densidade $\bar{\rho}$ pondo $\bar{\rho}(x) = \rho(x)$ se $x \in K$, $\bar{\rho}(x) = 0$ se $x \in M - K$, e escrevemos $\text{vol.}(K) = \int_M \bar{\rho}$, por definição.

Tem sentido também definir a integral de uma função contínua $f: M \to \mathbb{R}$ com suporte compacto. Basta tomar, por definição, $\int_M f = \int_W f \cdot \rho$, onde ρ é a densidade de volume de M.

7 Superfícies com bordo

Neste seção, ampliaremos o conceito de superfície, de modo que ele venha a incluir, por exemplo, as bolas fechadas no espaço euclidiano. Para isto, admitiremos que as parametrizações sejam definidas não apenas em subconjuntos abertos no espaço $\mathbb{R}^m$ mas possam ter abertos em semi-espaços como domínios.

Um *semi-espaço* num espaço vetorial E é um conjunto do tipo $H = \{x \in E; \alpha(x) \leq 0\}$, onde $\alpha \in E^*$ é um funcional linear não-nulo. O *bordo* do semi-espaço H é o conjunto $\partial H = \{x \in E; \alpha(x) = 0\}$. Evidentemente, ∂H é um subespaço vetorial de codimensão 1 em E.

Um semi-espaço $H \subset \mathbb{R}^m$ é reunião disjunta $H = \text{int.}\, H \cup \partial H$ do seu interior em $\mathbb{R}^m$ com o seu bordo. Os subconjuntos $A \subset H$, abertos com H, são de dois tipos: 1º) $A_1 \subset \text{int.}\, H$; neste caso, A_1 é também aberto em $\mathbb{R}^m$. 2º) $A_2 \cap \partial H \neq \varnothing$, então A_2 não é aberto em $\mathbb{R}^m$, pois nenhuma bola com centro num ponto $x \in \partial H$ pode estar contida em H.

Dois tipos de aberto num semi-espaço H.

Como vimos na seção 5, uma aplicação $f\colon X \to \mathbb{R}^n$, definida num subconjunto $X \subset \mathbb{R}^m$, diz-se diferenciável (respect. de classe C^k, $k \geq 1$) quando é a restrição de uma aplicação diferenciável (respect. de classe C^k) $F\colon U \to \mathbb{R}^n$, definida num aberto $U \subset \mathbb{R}^m$. Em particular, esta definição se aplica quando X é um subconjunto aberto de um semi-espaço, caso em que estaremos especialmente interessados.

Em geral, a derivada de uma aplicação diferenciável $f\colon X \to \mathbb{R}^n$ num ponto $x \in X$ não está bem definida, pois as possíveis extensões de f a vizinhanças de X podem ter diferentes derivadas no ponto x. [Isto é óbvio, por exemplo, no caso extremo em que X se reduz ao único ponto x.]

Mas se $A \subset H$ é aberto no semi-espaço $H \subset \mathbb{R}^m$ e $f\colon A \to \mathbb{R}^n$ é diferenciável então, para cada $x \in A$, a derivada $f'(x)\colon \mathbb{R}^m \to \mathbb{R}^m$ está bem definida.

Mais precisamente, como veremos agora, todas as extensões diferenciáveis $F\colon U \to \mathbb{R}^n$ de f a um aberto $U \subset \mathbb{R}^m$ contendo A possuem, no ponto $x \in A$, a mesma derivada $F'(x)\colon \mathbb{R}^m \to \mathbb{R}^n$, a qual indicaremos com $f'(x)$ e chamaremos a *derivada de f no ponto x.*

Isto é claro se $x \in \text{int.}\, H$. Vamos prová-lo quando $x \in A \cap \partial H$. Em primeiro lugar, observamos que existe uma base $\{v_1, \ldots, v_m\}$ em $\mathbb{R}^m$, tal que $v_1, \ldots, v_m \in H$. Com efeito, dado um vetor arbitrário $v \in \mathbb{R}^m$, temos $v \in H$ ou $-v \in H$. Assim, tomada uma base qualquer de $\mathbb{R}^m$ e trocando-se o sinal de cada um dos elementos que não pertencem a H, obtemos ainda uma base de $\mathbb{R}^m$ formada por vetores pertencentes a H. Segue-se que, para $x \in \partial H$ e $t \geq 0$ quaisquer, tem-se $x+t\cdot v_1 \in H, \ldots, x+t_m\cdot \in H$. De fato, se $H = \{y \in \mathbb{R}^m; \alpha(y) \leq 0\}$ então $\alpha(x+t\cdot v_i) = \alpha(x)+t\cdot\alpha(v_i) = t \cdot \alpha(v_i) \leq 0$ para $i = 1, \ldots, m$. Se, mais particularmente, tivermos $x \in A \cap \partial H$ então, para todo $t \geq 0$ suficientemente pequeno, vale ainda $x + t \cdot v_i \in A$ pois A é aberto em H.

Voltando à extensão diferenciável $F\colon U \to \mathbb{R}^n$ da aplicação f, sabemos que, para todo $v \in \mathbb{R}^m$, existe o limite $F'(x)\cdot v = \lim\limits_{v\to 0}[F(x+tv) - F(x)]/t$. Em particular, se fizermos $t \to 0$ por valores positivos, teremos ainda, para $i = 1,\ldots,m$:

$$F'(x)\cdot v_i = \lim_{t\to 0^+}\frac{F(x+tv_i)-F(x)}{t} = \lim_{t\to 0^+}\frac{f(x+tv_i)-f(x)}{t},$$

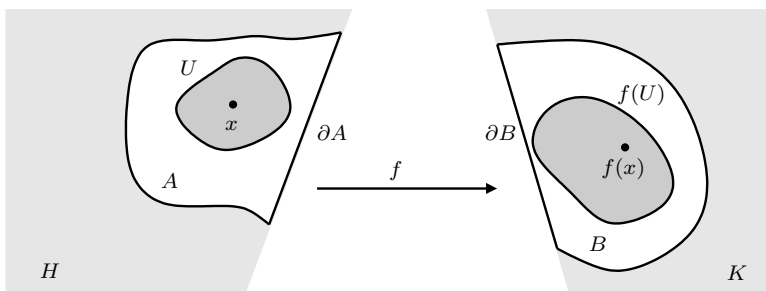
pois $x + tv_i \in A$ para todo $t \geq 0$ suficientemente pequeno. Assim, os valores da transformação linear $F'(x)\colon \mathbb{R}^m \to \mathbb{R}^n$ nos elementos básicos $v_1,\ldots,v_m$ são univocamente determinados a partir de f, logo a derivada $F'(x)$ não depende da extensão F escolhida.

Vale a Regra da Cadeia: se $f\colon A \to \mathbb{R}^n$ e $g\colon B \to \mathbb{R}^p$ são diferenciáveis em abertos de semi-espaços e $f(A) \subset B$ então $g\circ f\colon A \to \mathbb{R}^p$ é diferenciável, com $(g\circ f)'(x) = g'(f(x))\cdot f'(x)$. Isto decorre imediatamene da Regra da Cadeia usual, aplicada às extensões de f e g.

Seja $H \subset \mathbb{R}^m$ um semi-espaço. O *bordo* de um aberto $A \subset H$ é, por definição, o conjunto $\partial A = A \cap \partial H$. Observemos que ∂A é uma hiperfície em $\mathbb{R}^m$. [Com efeito, sendo A aberto em H, temos $A = U \cap H$, com $U \subset \mathbb{R}^m$ aberto. Então $U \cap \partial H = U \cap (H \cap \partial H) = (U\cap H)\cap \partial H = A \cap \partial H = \partial A$, logo ∂A é um subconjunto aberto da hiperfície $\partial H = \alpha^{-1}(0)$.]

O bordo de um aberto $A \subset H$ é invariante por difeomorfismos. Antes de demonstrar este fato, lembremos que um difeomorfismo (de classe C^k) entre dois abertos $A \subset H \subset \mathbb{R}^m$ e $B \subset K \subset \mathbb{R}^n$ de semi-espaços é uma bijeção diferenciável (de classe C^k) $f\colon A \to B$, cuja inversa $f^{-1}\colon B \to A$ também é diferenciável (de classe C^k). Pela Regra da Cadeia, das igualdades $f^{-1}\circ f = \mathrm{id}_A$ e $f\circ f^{-1} = \mathrm{id}_B$ concluimos que, para todo $x \in A$, com $y = f(x)$, $f'(x)\colon \mathbb{R}^m \to \mathbb{R}^n$ e $(f^{-1})'(y)\colon \mathbb{R}^n \to \mathbb{R}^m$ são isomorfismos, inversos um do outro. Em particular, $m = n$. A invariância de ∂A é expressa pelo

Teorema 13. *Sejam $A \subset H$, $B \subset K$ abertos em semi-espaços de $\mathbb{R}^m$. Se $f\colon A \to B$ é um difeomorfismo de classe C^1 então $f(\partial A) = \partial B$. Em particular, a restrição $f|\partial A$ é um difeomorfismo entre as hiperfícies ∂A e ∂B.*

Demonstração: Consideremos um ponto $x \in \mathrm{int.}\,A$, isto é, existe $U \subset \mathbb{R}^m$ aberto tal que $x \in U \subset A$. Restrito a U, f é um difeomorfismo de classe C^1 sobre sua imagem $f(U)$. Pelo Teorema da Aplicação Inversa, $f(U)$ é aberto em $\mathbb{R}^m$. Como $f(U) \subset B$,

segue-se que $f(x) \in \text{int.}\, B$. Isto significa que $f(\text{int.}\, A) \subset \text{int.}\, B$, logo $f^{-1}(\partial B) \subset \partial A$. Analogamente, $f(\partial A) \subset \partial B$. Logo $f(\partial A) = \partial B$.

Observação: Resulta do "Teorema da Invariância do Domínio", (demonstrado em Topologia) que o teorema acima vale, mais geralmente, quando f é apenas um homeomorfismo de A sobre B.

Estenderemos agora o conceito de parametrização.

Uma *parametrização* (de classe C^k e dimensão m) de um conjunto $U \subset \mathbb{R}^n$ é um homeomorfismo $\varphi \colon U_0 \to U$ de classe C^k, definido num aberto U_0 de um semi-espaço de $\mathbb{R}^m$, tal que $\varphi'(u) \colon \mathbb{R}^m \to \mathbb{R}^n$ é uma transformação linear injetiva, para cada $u \in U_0$.

Um conjunto $M \subset \mathbb{R}^n$ chama-se uma *superfície com bordo* (de dimensão m e classe C^k) quando cada ponto $x \in M$ pertence a um aberto $U \subset M$ que é imagem de uma parametrização $\varphi \colon U_0 \to U$, de classe C^k num aberto U_0 de algum semi-espaço de $\mathbb{R}^m$.

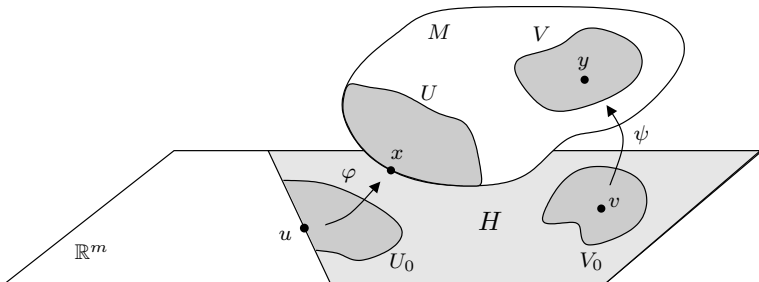
Examinemos as mudanças de parametrização numa superfície com bordo M, de classe C^k $(k \geq 1)$ e dimensão $m+1$, contida em $\mathbb{R}^n$.

Sejam $\varphi \colon U_0 \to U$, $\psi \colon V_0 \to V$ parametrizações de classe C^k de abertos $U, V \subset M$, com $U \cap V \neq \varnothing$. Como no caso sem bordo, a mudança de parametrização $\psi^{-1} \circ \varphi \colon \varphi^{-1}(U \cap V) \to \psi^{-1}(U \cap V)$ é um difeomorfismo de classe C^k. Para provar isto, basta mostrar que $\psi^{-1} \circ \varphi \in C^k$, pois $\varphi^{-1} \circ \psi = (\psi^{-1} \circ \varphi)^{-1}$ será também de classe C^k, pela mesma razão.

Tomemos um ponto qualquer $u \in \varphi^{-1}(U \cap V)$. Sejam $x = \varphi(u)$ e $v = \psi^{-1}(\varphi(u))$. Sabemos que ψ se estende a uma aplicação $\Psi \colon W \to \mathbb{R}^n$, de classe C^k num aberto W de $\mathbb{R}^m$, contendo v. Como $\Psi'(v)$ é injetiva, segue-se da Forma Local das Imersões que (restringindo W se necessário) Ψ é um homeomorfismo de W sobre sua imagem e o homeomorfismo inverso é a restrição a $\Psi(W)$ de uma aplicação Λ de classe C^k num aberto de $\mathbb{R}^n$. Logo, pondo $A = \varphi^{-1}(\Psi(W))$, vemos que $A \ni u$ e é um

aberto num semi-espaço de $\mathbb{R}^m$. Além disso, $(\psi^{-1} \circ \varphi)|A = (\Lambda \circ \varphi)|A$ é de classe C^k. Assim, $\psi^{-1} \circ \varphi$ é de classe C^k na vizinhança de cada ponto $u \in \varphi^{-1}(U \cap V)$, donde $\psi^{-1} \circ \varphi \in C^k$.

Seja M uma superfície com bordo. O *bordo* de M é o conjunto ∂M formado pelos pontos $x \in M$ tais que, para toda parametrização $\varphi\colon U_0 \to U$ de classe C^1 de um aberto $U \subset M$, com $x = \varphi(u)$, tem-se necessariamente $u \in \partial U_0$.

Pelo Teorema 13, juntamente com o fato de que cada mudança de parametrização é um difeomorfismo, dado $x \in M$, basta que exista uma parametrização $\varphi\colon U_0 \to U$ de classe C^1 de um aberto $U \subset M$, com $x = \varphi(u)$ e $u \in \partial U_0$, para que se tenha $x \in \partial M$. Dito de outro modo, se M é uma superfície com bordo e $x = \varphi(u) \in U$ para alguma parametrização $\varphi\colon U_0 \to U$, de classe C^1 num aberto $U_0 \subset \mathbb{R}^m$ então, para qualquer outra parametrização $\psi\colon V_0 \to V \subset M$, de classe C^1 num aberto V_0 de algum semi-espaço em $\mathbb{R}^m$, se $x = \psi(v)$, deve-se ter $v \in \operatorname{int.} V_0$.

Parametrizações numa superfície com bordo: $x \in \partial M$, $y \notin \partial M$

As superfícies definidas no Capítulo V serão chamadas "superfícies sem bordo". Elas correspondem, na definição acima, ao caso $\partial M = \varnothing$. Sempre que dissermos simplesmente "superfície", sem qualificações, estaremos nos referindo a uma superfície sem bordo.

Se M é uma superfície com bordo, de classe C^k e dimensão $m+1$, seu bordo ∂M é uma superfície (sem bordo) de classe C^k e dimensão m. As parametrizações que caracterizam ∂M como superfície são as restrições ao bordo $\partial U_0 = U_0 \cap \partial H$, das parametrizações $\varphi\colon U_0 \to U$, de classe C^k, que têm como imagem um aberto $U \subset M$ tal que $U \cap \partial M \neq \varnothing$.

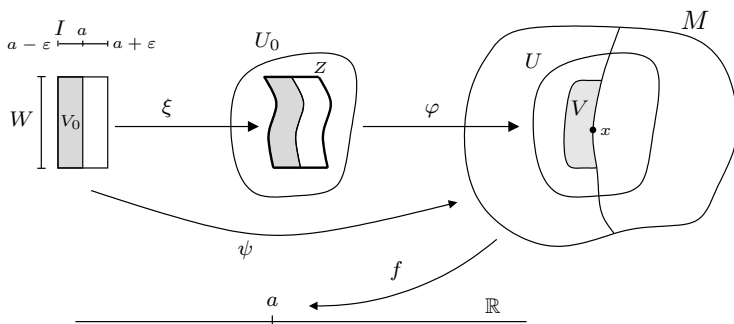
A restrição $\varphi|\partial U_0\colon \partial U_0 \to \partial U$ tem $\partial U = U \cap \partial M$ como imagem e seu domínio é o subconjunto aberto ∂U_0 do espaço vetorial m-dimensional

∂H. Para obter uma parametrização em ∂M que seja definida num aberto de $\mathbb{R}^m$, devemos tomar uma base em ∂H e representar cada elemento $u \in \partial H$ pela lista de suas m coordenadas em relação a tal base.

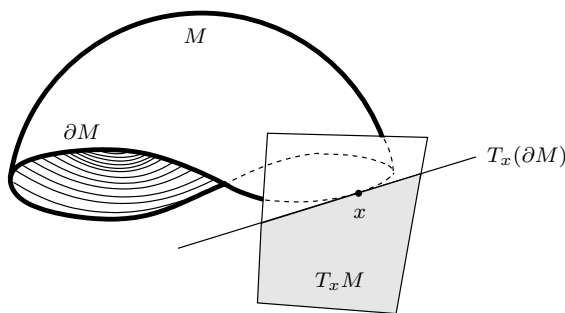
Outra maneira de parametrizar ∂U é a seguinte. Escrevemos os elementos de $\mathbb{R}^{m+1}$ sob a forma $u = (u_0, u_1, \ldots, u_m)$, pomos $H_0 = \{u \in \mathbb{R}^{m+1}; u_0 \leq 0\}$, identificamos ∂H_0 com $\mathbb{R}^m$ pela correspondência $(0, v_1, \ldots, v_m) \mapsto (v_1, \ldots, v_m)$ e "padronizamos"as parametrizações de classe C^k em M, considerando apenas aquelas que são definidas em subconjuntos abertos do semi-espaço H_0. Para todo semi-espaço $H \subset \mathbb{R}^{m+1}$ existe um isomorfsmo linear $T\colon \mathbb{R}^{m+1} \to \mathbb{R}^{m+1}$ tal que $T(H_0) = H$. Então, dada uma parametrização $\psi\colon V_0 \to U$, de classe C^k no aberto $V_0 \subset H$, pomos $U_0 = T^{-1}(V_0)$ e obtemos $\varphi = \psi \circ T\colon U_0 \to U$, uma parametrização *padronizada* (isto é, definida num aberto de H_0), de classe C^k e com mesma imagem que ψ. Se $\varphi\colon U_0 \to U$ é padronizada e $U \cap \partial M \neq \varnothing$, a restrição $\varphi|\partial U_0\colon \partial U_0 \to \partial U$ é uma parametrização na superfície ∂M, definida num subconjunto aberto $\partial U_0 \subset \mathbb{R}^m$.

O teorema seguinte é fonte de exemplos de superfícies com bordo.

Teorema 14. *Seja $f\colon M \to \mathbb{R}$ uma função real de classe C^k numa superfície M, de classe C^k e dimensão $m+1$. Se $a \in \mathbb{R}$ é valor regular de f então o conjunto $N = \{x \in M; f(x) \leq a\}$ é uma superfície de classe C^k, com dimensão $m+1$, cujo bordo é $\partial N = f^{-1}(a)$.*

Demonstração: Evidentemente, o conjunto $A = \{x \in M; f(x) < a\}$ é aberto em M, logo é uma superfície de dimensão $m+1$ e classe C^k. Resta então parametrizarmos as vizinhanças dos pontos $x \in N$ tais que $f(x) = a$. Dado x, seja $\varphi\colon U_0 \to U$ uma parametrização de classe C^k de um aberto U em M, tal que $x = \varphi(u) \in U$, com $u = (u_0, u_1, \ldots, u_m)$. Como a é valor regular de f, e portanto da função $f \circ \varphi\colon U_0 \to \mathbb{R}$, podemos, sem perda de generalidade, supor que $\dfrac{\partial(f \circ \varphi)}{\partial u_m}(u) > 0$. Pela Forma Local das Submersões, existem um aberto $W \subset \mathbb{R}^m$ contendo $(u_1, \ldots, u_m)$, um intervalo $I = (a-\varepsilon, a+\varepsilon)$ e um difeomorfismo $\xi\colon W \times I \to Z$ de classe C^k sobre um aberto $Z \subset U_0$, tal que $f \circ \varphi \circ \xi\colon W \times I \to \mathbb{R}$ tem a forma $f \circ \varphi \circ \xi\colon (z, t) \mapsto t$. Consideremos em $\mathbb{R}^{m+1}$ o semi-espaço H, formado pelos pontos cuja última coordenada é $\leq a$. Ponhamos $V_0 = (W \times I) \cap H$, $\psi = (\varphi \circ \xi)|V_0$, $V = \varphi(\xi(V_0))$. Então $\psi\colon V_0 \to V$ é uma parametrização do aberto $V \subset N$, com $x \in V$.

Por exemplo, a bola fechada de centro num ponto $x_0 \in \mathbb{R}^{m+1}$ e raio $a > 0$ é o conjunto dos pontos $x \in \mathbb{R}^{m+1}$ tais que $f(x) \leq a^2$, onde $f\colon \mathbb{R}^{m+1} \to \mathbb{R}$ é definida por $f(x) = |x - x_0|^2 = \langle x - x_0, x - x_0 \rangle$. Como o único valor não-regular de f é 0, segue-se que a bola fechada é uma superfície, cujo bordo é a esfera.

Seja $M \subset \mathbb{R}^n$ uma superfície com bordo, de classe C^1 e dimensão $m+1$. A cada ponto $x \in M$ (mesmo os pontos do bordo!) associaremos um subespaço vetorial $T_xM \subset \mathbb{R}^n$, de dimensão $m+1$, chamado o *espaço vetorial tangente* a M no ponto x, o qual é definido (analogamente ao caso $\partial M = \varnothing$) como a imagem $\varphi'(u) \cdot \mathbb{R}^{m+1}$, onde $\varphi\colon U_0 \to U$ é qualquer parametrização de classe C^1 de um aberto $U \subset M$, com $x = \varphi(u) \in U$.

Se $x \in \partial M$, então U_0 é aberto num semi-espaço $H \subset \mathbb{R}^{m+1}$, com $u = \varphi^{-1}(x) \in \partial H$. A imagem $\varphi'(u) \cdot \partial H = T_x(\partial M)$ é o espaço vetorial tangente ao bordo ∂M no ponto x. Evidentemente, $T_x(\partial M) \subset T_xM$ é um subespaço vetorial de codimensão 1.

A definição de T_xM faz uso da parametrização $\varphi\colon U_0 \to U$, com $x = \varphi(u) \in U$. Mas se $\psi\colon V_0 \to V$ é outra parametrização C^1 de um aberto $V \subset M$, com $x = \psi(v) \in V$, então $\xi = \psi^{-1} \circ \varphi\colon \varphi^{-1}(U \cap V) \to \psi^{-1}(U \cap$

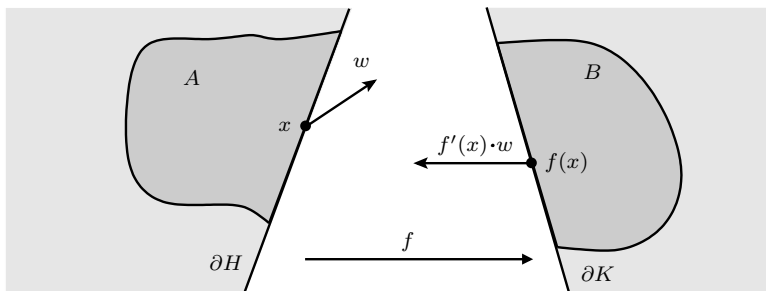
V) é um difeomorfismo, com $\psi \circ \xi = \varphi$, logo $\psi'(v) \cdot \xi'(u) = \varphi'(u)$. Como $\xi'(u)\colon \mathbb{R}^{m+1} \to \mathbb{R}^{m+1}$ é um isomorfismo, segue-se que $\varphi'(u) \cdot \mathbb{R}^{m+1} = \psi'(v) \cdot \mathbb{R}^{m+1}$. Portanto o espaço tangente $T_xM = \varphi'(u) \cdot \mathbb{R}^{m+1}$ independe da parametrização usada para definí-lo.

Se $x \in \partial M$, no espaço vetorial tangente T_xM não somente temos um subespaço vetorial distinguido $T_x(\partial M)$ como temos ainda um semi-espaço, formado pelos vetores tangentes ao bordo mais os vetores que apontam para fora da superfície M. (Região pontilhada na figura acima.) Para dar sentido preciso a essa noção, começaremos com um semi-espaço $H \subset \mathbb{R}^m$.

Diz-se que o vetor $w \in \mathbb{R}^m$ *aponta para fora* do semi-espaço $H \subset \mathbb{R}^m$ quando $w \notin H$. Se H é definido pela desigualdade $\alpha(x) \leq 0$, isto significa que $\alpha(w) > 0$. Este conceito é invariante por difeomorfismos, conforme o

Teorema 15. *Seja $f\colon A \to B$ um difeomorfismo entre abertos $A \subset H$, $B \subset K$ de semi-espaços $H, K \subset \mathbb{R}^m$. Se $w \in \mathbb{R}^m$ aponta para fora de H então, para cada $x \in \partial A$, o vetor $f'(x) \cdot w$ aponta para fora de K.*

Demonstração: Sejam $\alpha(x) \leq 0$ e $\beta(x) \leq 0$ respectivamente as desigualdades que definem os semi-espaços H e K. Como, pelo Teorema 13, f é um difeomorfismo entre ∂A e ∂B, a derivada $f'(x)\colon \mathbb{R}^m \to \mathbb{R}^m$ transforma ∂H isomorficamente sobre ∂K. Assim, dado $v \in \mathbb{R}^m$, tem-se $\beta[f'(x) \cdot v] = 0$ se, e somente se,

$\alpha(v) = 0$. Como $\alpha(w) > 0$, basta provarmos que $\beta[f'(x) \cdot w] \geq 0$. Ora, se $t < 0$ então $x + t \cdot w \in H$, logo, para todo t negativo e suficientemente próximo de zero, temos $x + t \cdot w \in A - \partial A$, donde $f(x + t \cdot w) \in B - \partial B$ e $\beta f(x + t \cdot w) < 0$. Para tais valores de t, então $[\beta f(x + t \cdot w) - \beta f(x)]/t = \beta f(x + tw)/t > 0$. Passando ao limite quando $t \to 0^-$, obtemos $\beta[f'(x) \cdot w] \geq 0$, o que demonstra o teorema.

Dada a superfície com bordo $M \subset \mathbb{R}^n$, seja $x \in \partial M$. Diz-se que um vetor $w \in T_xM$ *aponta para fora* da superfície M quando existe uma parametrização $\varphi\colon U_0 \to U$, de classe C^1 no aberto U_0 de um semi-espaço $H \subset \mathbb{R}^{m+1}$, com valores no aberto $U \subset M$, tal que $x = \varphi(u) \in U$ e $w = \varphi'(u) \cdot w_0$, onde $w_0 \in \mathbb{R}^{m+1}$ aponta para fora do semi-espaço H. Neste caso, segue-se do Teorema 15 que, para qualquer outra parametrização $\psi\colon V_0 \to V$, com $x = \psi(v) \in V$, tem-se $w = \psi'(v) \cdot w_1$, onde $w_1 \in \mathbb{R}^{m+1}$ aponta para fora do semi-espaço $K \subset \mathbb{R}^{m+1}$, no qual V_0 é aberto. [Com efeito, pondo $\xi = \psi^{-1} \circ \varphi$ e $w_1 = \xi'(u) \cdot w_0$, temos $\psi'(v) \cdot w_1 = \psi'(v) \cdot \xi'(u) \cdot w_0 = \varphi'(u) \cdot w_0 = w$ e, pelo Teorema 15, w_1 aponta para fora de K pois w_0 aponta para fora de H.]

Em cada ponto $x \in \partial M$, os vetores tangentes a ∂M juntamente com os vetores que apontam para fora de M formam um semi-espaço de T_xM. Entre os vetores que apontam para fora de M no ponto x, existe um único que tem comprimento 1 e é normal a $T_x(\partial M)$. Indicando-o com $v(x)$, obtemos um campo de vetores unitários $v\colon \partial M \to \mathbb{R}^n$, normais a ∂M. Se M é de classe C^k, o campo v é de classe C^{k-1}. Com efeito, se $\varphi\colon U_0 \to U$ é uma parametrização de classe C^k definida no aberto U_0 do semi-espaço $H \subset \mathbb{R}^{m+1}$, tomamos uma base positiva $\{v_0, v_1, \ldots, v_m\} \subset \mathbb{R}^{m+1}$ tal que v_0 aponta para fora de H e $\{v_1, \ldots, v_m\} \subset \partial H$. Então

$$v(x) = \frac{\varphi'(u) \cdot v_1 \times \cdots \times \varphi'(u) \cdot v_m}{|\varphi'(u) \cdot v_1 \times \cdots \times \varphi'(u) \cdot v_m|}$$

para todo $x = \varphi(u) \in \partial U = \partial M \cap U$. [Aqui, estamos considerando o produto vetorial de m vetores no espaço T_xM, de dimensão $m+1$, orientado pela condição de que a base $\{\varphi'(u) \cdot v_0, \varphi(u) \cdot v_1, \ldots, \varphi'(u) \cdot v_m\}$ seja positiva.] Resulta daí que se M é uma superfície com bordo, de classe C^1 e dimensão $m+1$ no espaço $\mathbb{R}^{m+1}$ então seu bordo ∂M é uma hiperfície orientável. (Cfr. Teorema 15, Capítulo V.)

Uma superfície com bordo M diz-se *orientável* quando admite um atlas coerente de classe C^1. Como no caso sem bordo, o atlas $\mathcal{A}$ chama-se *coerente* quando, dadas $\varphi, \psi \in \mathcal{A}$ quaisquer, a mudança de parametrização $\psi^{-1} \circ \varphi$ tem determinante jacobiano positivo em todos os pontos do seu domínio. [Evidentemente, para que $\psi^{-1} \circ \varphi$ faça sentido, as imagens de φ e ψ devem ter pontos em comum.]

Mostraremos agora que se M é orientável então seu bordo ∂M também o é. Isto é evidente quando dim. $M = 1$ pois ∂M tem então dimensão zero. Seja pois M *orientada*, com $\dim M = m+1$, $m > 0$.

Chamemos de $\mathcal{A}$ o conjunto das parametrizações $\varphi\colon U_0 \to U$ em M, de classe C^1, com as seguintes propriedades:

0) U_0 é conexo;

1) U_0 é aberto no semi-espaço $H_0 = \{(u_0, \ldots, u_m) \in \mathbb{R}^{m+1}; u_0 \leq 0\}$;

2) φ é positiva relativamente à orientação de M.

Já vimos que o conjunto das parametrizações $\varphi\colon U_0 \to U$ em M que cumprem a condição 1) (isto é, são "padronizadas") constituem um atlas em M. Se impusermos, ademais, a condição 2), teremos ainda um atlas. Com efeito, dada $\psi\colon V_0 \to U$, cumprindo 0) e 1), se ψ não for positiva podemos compô-la com a transformação linear $T\colon (u_0, \ldots, u_m) \mapsto (u_0, \ldots, -u_m)$ e, pondo $U_0 = T^{-1}(V_0)$, veremos que $\varphi = \psi \circ T\colon U_0 \to U$ cumpre a condição 0), a condição 1) (porque sendo $m > 0$, T leva H_0 em si mesmo) e é positiva, pois det. $T < 0$. Assim, $\mathcal{A}$ é um atlas em M.

Identifiquemos, como antes, $\mathbb{R}^m$ com $\partial H_0 = \{u \in \mathbb{R}^{m+1}; u_0 = 0\}$.

Seja $\mathcal{A}_0$ o conjunto das restrições $\varphi_0 = \varphi|\partial U_0$, das parametrizações $\varphi \in \mathcal{A}$ tais que $\partial U_0 = U_0 \cap \mathbb{R}^m \neq \varnothing$.

Evidentemente, $\mathcal{A}_0$ é um atlas de classe C^1 em ∂M.

Afirmamos que o atlas $\mathcal{A}_0$ é coerente. Com efeito, se $\varphi_0\colon \partial U_0 \to \partial U$ e $\psi_0\colon \partial V_0 \to \partial V$ pertencem a $\mathcal{A}_0$, com $\partial U \cap \partial V \neq \varnothing$ então a mudança de parametrização $\xi_0 = \psi_0^{-1} \circ \varphi_0$ é a restrição do difeomorfismo $\xi = \psi^{-1} \circ \varphi$ ao bordo do seu domínio. Seja $A = \xi'(u)\colon \mathbb{R}^{m+1} \to \mathbb{R}^{m+1}$ a derivada de ξ num ponto qualquer u do domínio de ξ_0. [Isto é, $\varphi(u) \in \partial U \cap \partial V$.] Como $\mathcal{A}$ é coerente, det. $A > 0$. Como ξ é um difeomorfismo do aberto $\varphi^{-1}(U \cap V) \subset H_0$ no aberto $\psi^{-1}(U \cap V) \subset H_0$ segue-se do Teorema 13 que $A \cdot (\partial H_0) = \partial H_0$, ou seja $A \cdot e_i = (0, a_{1i}, \ldots, a_{mi})$ para todo $i = 1, \ldots, m$. Finalmente como $e_0 = (1, 0, \ldots, 0)$ aponta para fora de H_0, segue-se do Teorema 15 que $A \cdot e_0 = (a_{00}, a_{10}, \ldots, a_{m0})$ também aponta para fora de H_0, isto é, $a_{00} > 0$. Assim, a matriz de A tem a forma

$$\begin{pmatrix} a_{00} & 0 & \ldots & 0 \\ a_{10} & a_{11} & \ldots & a_{1m} \\ \vdots & & & \\ a_{m0} & a_{m1} & \ldots & a_{mm} \end{pmatrix},$$

com $a_{00} > 0$. Segue-se que det. $A = a_{00} \cdot$ det. A_0, onde $A_0 = A|\mathbb{R}^m$ é a derivada de ξ_0 no ponto u. (A matriz de A_0 é formada pelos elementos

a_{ij} com $i > 0$ e $j > 0$.) Vemos portanto que det. $A_0 > 0$, logo $\mathcal{A}_0$ é coerente.

A orientação definida pelo atlas $\mathcal{A}_0$ em ∂M diz-se *induzida* pela orientação de M.

Relativamente à orientação induzida por M em ∂M, uma base $\{w_1, \ldots, w_m\} \subset T_x(\partial M)$ é positiva se, e somente se, para qualquer vetor $w_0 \in T_xM$, que aponte para fora de M, $\{w_0, w_1, \ldots, w_m\}$ é uma base positiva de T_xM.

Em particular, se $v(x) \in T_xM$ é o vetor unitário, tangente a M e normal a ∂M no ponto x, que aponta para fora de M, então $\{w_1, \ldots, w_m\} \subset T_x(\partial M)$ é uma base positiva se, e somente se, a base $\{v(x), w_1, \ldots, w_m\} \subset T_xM$ é positiva.

Com efeito, uma base $\{w_1, \ldots, w_m\} \subset T_x(\partial M)$ é positiva se, e somente se,

$$w_j = \sum_{i=1}^{m} a_{ij} \frac{\partial \varphi}{\partial u_i}(u), \quad j = 1, \ldots, m,$$

onde a matriz $A_0 = (a_{ij})$, com m linhas e m colunas, tem determinante positivo, sendo $\varphi \colon U_0 \to U$ uma parametrização definida num aberto U_0 do semi-espaço $H_0 = \{(u_0, u_1, \ldots, u_m) \in \mathbb{R}^{m+1}; u_0 \leq 0\}$, com $u \in \partial H_0$, $\varphi(u) = x$ e φ positiva (relativamente à orientação de M).

Como $e_0 = (1, 0, \ldots, 0)$ aponta para fora de H_0, vê-se que

$$\frac{\partial \varphi}{\partial u_0}(u) = \varphi'(u) \cdot e_0 \in T_xM$$

aponta para fora da superfície M. Portanto, se $w_0 \in T_xM$ é qualquer vetor que aponte para fora de M, temos

$$w_0 = a_{00} \frac{\partial \varphi}{\partial u_0}(u) + a_{10} \frac{\partial \varphi}{\partial u_1}(u) + \cdots + a_{m0} \frac{\partial \varphi}{\partial u_m}(u), \text{ com } a_{00} > 0.$$

Para $j = 1, \ldots, m$, temos

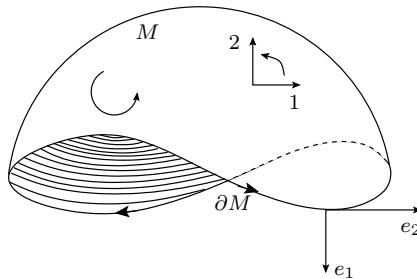
$$w_j = 0 \cdot \frac{\partial \varphi}{\partial u_0}(u) + a_{1j} \frac{\partial \varphi}{\partial u_1}(u) + \cdots + a_{mj} \frac{\partial \varphi}{\partial u_m}(u).$$

Assim, a matriz A, de passagem da base

$$\left\{ \frac{\partial \varphi}{\partial u_0}(u), \frac{\partial \varphi}{\partial u_1}(u), \ldots, \frac{\partial \varphi}{\partial u_m}(u) \right\}$$

para a base $\{w_0, w_1, \ldots, w_m\}$, tem a forma

$$A = \begin{pmatrix} a_{00} & 0 & \cdots & 0 \\ a_{10} & & & \\ \vdots & & A_0 & \\ a_{m0} & & & \end{pmatrix}$$

Logo det. $A = a_{00} \cdot$ det. A_0, ou seja, det. $A > 0 \Leftrightarrow$ det. $A_0 > 0$. Isto significa que, quando $w_0 \in T_xM$ aponta para fora de M, $\{w_1, \ldots, w_m\} \subset T_x(\partial M)$ é uma base positiva se, e somente se, $\{w_0, w_1, \ldots, w_m\} \subset T_xM$ é positiva.

A orientação induzida no bordo.

Observação: O intervalo $[0,1]$ é uma "superfície com bordo", de dimensão 1 e classe C^∞, munida do atlas $\mathcal{A} = \{\varphi, \psi\}$, onde $\varphi\colon [0,1) \to [0,1)$ e $\psi\colon (0,1] \to (0,1]$ são iguais à função identidade. O domínio de φ é um aberto da semi-reta $[0,+\infty)$, que é um semi-espaço de $\mathbb{R}$, e o domínio de ψ é um aberto no semi-espaço $(-\infty, 1] \subset \mathbb{R}$. A mudança de parametrização $\psi^{-1} \circ \varphi\colon (0,1) \to (0,1)$ também é a função identidade, cuja derivada (determinante jacobiano) é 1 em todos os pontos. Logo $\mathcal{A}$ é um atlas coerente, que define a orientação natural de $[0,1]$. Se quisermos padronizar as parametrizações em $[0,1]$, isto é, considerar apenas as que têm como domínio um aberto em $H_0 = (-\infty, 0]$, o modo mais simples é

tomar o atlas $\beta = \{\xi, \zeta\}$, onde as parametrizações $\xi\colon (-1,0] \to [0,1)$ e $\zeta\colon (-1,0] \to (0,1]$ são dadas por $\xi(t) = -t$ e $\zeta(t) = 1+t$. A mudança de parametrização $\zeta^{-1} \circ \xi\colon (-1,0) \to (-1,0)$ é a função $t \to -1-t$, cuja derivada é -1. Assim, o atlas β não é coerente. Na realidade, nenhum atlas em $[0,1]$ formado por parametrizações padronizadas pode ser coerente, pois nele qualquer vizinhança de 0 será parametrizada por uma função com derivada negativa, enquanto toda parametrização padronizada de uma vizinhança de 1 deverá ter derivada positiva. Por conseguinte, a definição de orientação induzida no bordo, dada acima para superfícies de dimensão ≥ 2, não se adapta para dimensão 1. Quando dim. $M = 1$, tem-se dim. $\partial M = 0$, ou seja, ∂M é um conjunto de pontos isolados. Orientar uma superfície de dimensão zero é, por definição, atribuir um sinal, $+$ ou $-$, a cada um dos seus pontos. Se dim. $M = 1$ e M está orientada, a orientação induzida no ponto $x \in \partial M$ será $+x$ se cada vetor que forma uma base positiva de T_xM apontar para fora de M e será $-x$ no caso contrário. Por exemplo, orientando $[0,1]$ por meio do atlas $\mathcal{A} = \{\varphi, \psi\}$ acima definido, a orientação induzida no bordo $\partial[0,1] = \{0,1\}$ atribui o sinal $+$ ao ponto 1 e o sinal $-$ ao ponto 0. Escrevemos, então $\partial[1,0] = \{+1\} \cup \{-0\}$. Analogamente, se considerarmos as semi-retas $M = (-\infty, a]$ e $N = [b, +\infty)$ orientadas pelos atlas, cada um dos quais consiste na aplicação identidade, teremos $\partial M = \{+a\}$ e $\partial N = \{-b\}$.

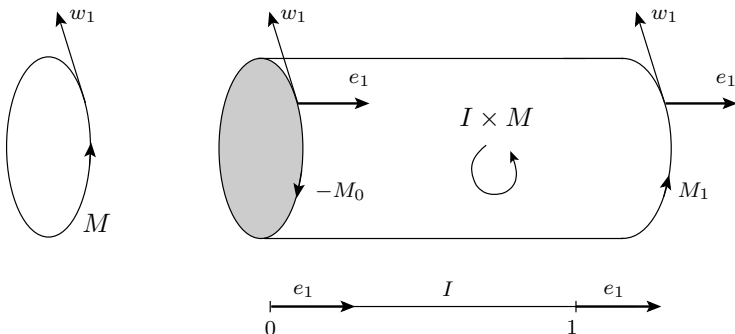
O produto cartesiano $M \times N$ de duas superfícies de classe C^k sem bordo é, como sabemos, uma superfície de classe C^k sem bordo, com um atlas formado pelas parametrizações $\varphi \times \psi\colon U_0 \times V_0 \to U \times V$, onde $(\varphi \times \psi)(u,v) = (\varphi(u), \psi(v))$. Se M e N são orientadas, as parametrizações $\varphi \times \psi$ tais que $\varphi > 0$ em M e $\psi > 0$ em N constituem um atlas coerente, que define em $M \times N$ a *orientação-produto*. Em cada ponto $(x,y) \in M \times N$, tem-se $T_{(x,y)}(M \times N) = T_xM \times T_yN$. Poderemos considerar T_xM e T_yN como subespaços vetoriais de $T_{(x,y)}(M \times N)$, se identificarmos M com $M \times y$ e N com $x \times N$. Sejam $\{v_1, \ldots, v_m\} \subset T_xM$ e $\{w_1, \ldots, w_n\} \subset T_yN$ bases positivas. A orientação do espaço tangente a $M \times N$ no ponto (x,y) é determinada pela exigência de que $\{v_1, \ldots, v_m, w_1, \ldots, w_n\} \subset T_{(x,y)}(M \times N)$ seja uma base positiva.

Quando ambas as superfícies M e N possuem bordo, o produto cartesiano $M \times N$ não é uma superfície com bordo. Isto se deve ao fato de que o produto cartesiano de dois semi-espaços não é (sequer difeomorfo a) um semi-espaço. Por exemplo, $[0,+\infty) \times [0,+\infty) = \{(x,y) \in$

$\mathbb{R}^2; x \geq 0, y \geq 0\}$ possui o "ponto anguloso" $(0,0)$. Também o quadrado $[0,1] \times [0,1]$ não é uma superfície com bordo: possui quatro pontos angulosos, nos quais o bordo não tem espaço vetorial tangente.

Mas se apenas a superfície M tem bordo, o produto cartesiano $M \times N$ é uma superfície com bordo e $\partial(M \times N) = \partial M \times N$, como se vê usando as parametrizações $\varphi \times \psi$. [Ocorre que se U_0 é aberto no semi-espaço $H \subset \mathbb{R}^m$ e V_0 é aberto em $\mathbb{R}^n$ então $U_0 \times V_0$ é aberto no semi-espaço $H \times \mathbb{R}^n \subset \mathbb{R}^{m+n}$, com $\partial(U_0 \times V_0) = \partial U_0 \times V_0$.]

Se a superfície com bordo M e a superfície sem bordo N são ambas orientáveis, então o produto cartesiano $M \times N$ é uma superfície orientável (com bordo) e a orientação produto é definida, como no caso em que $\partial M = \partial N = \varnothing$, pelo atlas $\mathcal{P}$, conjunto das parametrizações $\varphi \times \psi$, onde $\varphi > 0$ em M e $\psi > 0$ em N. A coerência de $\mathcal{P}$ decorre, como no Capítulo V, §14, da igualdade $(\varphi_2 \times \psi_2)^{-1} \circ (\varphi_1 \times \psi_1) = (\varphi_2^{-1} \circ \varphi_1) \times (\psi_2^{-1} \circ \psi_1)$. Evidentemente, o atlas coerente $\mathcal{P}$ não é máximo, logo há outras parametrizações positivas em $M \times N$. Por exemplo, a própria igualdade acima mostra que se $\varphi < 0$ em M e $\psi < 0$ em N então $\varphi \times \psi$ é compatível com toda parametrização de $\mathcal{P}$, logo $\varphi \times \psi$ é uma parametrização positiva em $M \times N$, isto é, pertence ao atlas coerente máximo que contém $\mathcal{P}$.

Estamos especialmente interessados no produto cartesiano $I \times M$ do intervalo $I = [0,1]$ pela superfície *orientada*, sem bordo, M. Consideremos I orientado como na Observação acima. Temos $\partial(I \times M) = M_0 \cup M_1$, onde $V_0 = \{0\} \times M$ e $M_1 = \{1\} \times M$. Tomemos em M_0 e M_1 , respectivamente, as orientações segundo as quais os difeomorfismos canônicos $f_0\colon M \to M_0$ e $f_1\colon M \to M_1$, dados por $f_0(x) = (0,x)$ e

$f_1(x) = (1, x)$, são positivos, isto é, preservam a orientação.

Afirmamos que a orientação induzida no bordo $\partial(I \times M) = M_0 \cup M_1$ pela orientação-produto de $I \times M$ coincide com a de M_1 e é oposta à de M_0. Se indicarmos com $-M_0$ a superfície M_0 munida com a orientação oposta da que definimos acima, poderemos escrever $\partial(I \times M) = M_1 \cup (-M_0)$.

Isto se prova observando que se $\varphi, \psi\colon U_0 \to U$ são parametrizações em M, com $\varphi > 0$ e $\psi < 0$ então $\Phi\colon (-1, 0] \times U_0 \to (0, 1] \times U$ e $\Psi\colon (-1, 0] \times U_0 \to [0, 1) \times U$, dadas por $\Phi(t, u) = (1+t, \varphi(u))$ e $\Psi(t, u) = (-t, \psi(u))$, são parametrizações positivas, padronizadas, em $I \times M$. A orientação induzida em M_1 é determinada pelo atlas $\mathcal{P}_1$, conjunto das restrições do tipo $\Phi|(\{0\} \times U_0)$, ou seja, $u \mapsto (1, \varphi(u))$, $\varphi > 0$. Esta é a orientação própria de M_1. A orientação induzida por $I \times M$ em M_0 é dada pelo atlas $\mathcal{P}_0$, conjunto das restrições do tipo $\Psi|(\{0\} \times U_0)$, ou seja, $u \mapsto (0, \psi(u))$. Como as parametrizações ψ são negativas, esta é a orientação oposta de M_0.

Outra maneira de chegar à mesma conclusão consiste em notar que, em cada ponto $(t, x) \in I \times M$, uma base positiva de $T_{(t,x)}(I \times M) = \mathbb{R} \times M$, relativamente à orientação-produto, tem a forma $\{e_1, w_1, \ldots, w_m\}$, onde e_1 é o vetor da base canônica de $\mathbb{R}$ (igual a 1!) e $\{w_1, \ldots, w_m\} \subset T_xM$ é uma base positiva na orientação de M. No ponto $t = 1$, o vetor e_1 aponta para fora de I, logo $e_1 \in T_{(1,x)}(I \times M)$ aponta para fora de $I \times M$ no ponto $(1, x)$. Sendo $\{e_1, w_1, \ldots, w_m\}$ uma base positiva de $T_{(1,x)}(I \times M)$, resulta que $\{w_1, \ldots, w_m\}$ é uma base positiva na orientação induzida em $\partial(I \times M)$ no ponto $(1, x)$, ou seja, em M_1. Por outro lado, e_1 aponta para dentro de I no ponto $t = 0$ e daí se segue que a mesma base $\{w_1, \ldots, w_m\}$ é negativa no ponto $(0, x) \in M_0$.

Formas diferenciais, diferencial exterior e integração de formas (ou de densidades escalares) se estudam em superfícies com bordo, sem nenhuma diferença essencial do caso $\partial M = \varnothing$. Faremos apenas um comentário a propósito de partições da unidade: elas existem em qualquer subconjunto $X \subset \mathbb{R}^m$. Mais precisamente, toda cobertura aberta $X = \cup V_\lambda$ é do tipo $V_\lambda = U_\lambda \cap X$ onde $X \subset \cup U_\lambda$ e cada U_λ é aberto em $\mathbb{R}^m$. A reunião $U = \cup U_\lambda$ é um aberto de $\mathbb{R}^m$, contendo X. Em particular, U é uma superfície sem bordo. Portanto existe uma partição da unidade $\Sigma\, \zeta_\lambda = 1$, de classe C^∞, estritamente subordinada à cobertura $U = \cup U_\lambda$. Pondo $\xi_\lambda = \zeta_\lambda | X$, obtemos uma partição da unidade de classe C^∞, $\Sigma\, \xi_\lambda = 1$ em X, estritamente subordinada à cobertura $X = \cup V_\lambda$, dada inicial-

mente. Se X é uma superfície de classe C^k (com ou sem bordo), as funções $\xi_\lambda\colon X \to \mathbb{R}$, embora sejam restrições de funções C^∞ definidas numa vizinhança aberta de X em $\mathbb{R}^m$, devem ser consideradas como funções de classe C^k em X (conforme observamos no parágrafo 5) pois esta é a classe que obtemos quando as compomos com parametrizações (de classe C^k) da superfície X.

8 O Teorema de Stokes

Demonstraremos neste parágrafo o teorema que contém os resultados clássicos de Análise Vetorial, vagamente enunciados no início da seção 3, conhecidos sob os nomes de Green, Riemann, Gauss, Ostrogradski e Stokes. Os parágrafos seguintes conterão aplicações desse teorema geral, hoje conhecido como Teorema de Stokes. [Na verdade, nenhuma de suas versões se deve a Stokes. Até mesmo o caso particular que em Análise Vetorial leva seu nome (caso de superfície com bordo em $\mathbb{R}^3$) foi descoberto por Lord Kelvin e apresentado por Stokes como questão num exame.]

Vejamos inicialmente um caso particular.

Teorema 16. *Numa superfície M, com bordo, orientada, de dimensão $m+1$, seja ω uma forma de classe C^1 e grau m, com suporte compacto, disjunto do bordo ∂M. Nestas condições $\int_M d\omega = 0$.*

Demonstração: Se $\omega = \omega_1 + \cdots + \omega_k$ então $\int_M d\omega = \sum_{i=1}^{k} \int_M d\omega_i$. Podemos supor então que o suporte de ω está contido na imagem de uma parametrização positiva $\varphi\colon U_0 \to U$, onde $U \cap \partial M = \varnothing$. O caso geral reduz-se a este mediante uma partição da unidade. Para todo $x = \varphi(u) \in U$, escrevamos

$$\omega(x) = \sum_{i=1}^{m+1} (-1)^{i+1}\, a_i(u) du_1 \wedge \cdots \wedge \widehat{du_1} \wedge \cdots \wedge du_{m+1}\,.$$

Então

$$d\omega(x) = \left[\sum_{i=1}^{m+1} \frac{\partial a_i}{\partial u_i}(u)\right] du_1 \wedge \cdots \wedge du_{m+1}\,.$$

Seja $K = \prod_{i=1}^{m+1} [\alpha_j, \beta_i]$ um bloco em $\mathbb{R}^{m+1}$, contendo $\varphi^{-1}(\text{supp.}\,\omega)$ Consideremos as funções $a_1, \ldots, a_{m+1}$ definidas (e contínuas) em K, pondo-as

iguais a zero em $K - U_0$. Em particular, cada a_i se anula na fronteira de K, isto é, $a_i(u_1, \ldots, u_{m+1}) = 0$ se algum $u_j = \alpha_j$ ou $u_j = \beta_j$. Para cada $i = 1, \ldots, m+1$, seja K_i o produto cartesiano dos intervalos $[\alpha_j, \beta_j]$, com exceção do i-ésimo. Temos:

$$\int_M d\omega = \sum_{i=1}^{m+1} \int_K \frac{\partial a_i}{\partial u_i}(u) du_1 \ldots du_{m+1} =$$

$$= \sum_{i=1}^{m+1} \int_{K_i} \left[\int_{\alpha_i}^{\beta_i} \frac{\partial a_i}{\partial u_i}(u) du_i \right] du_1 \ldots \widehat{du_i} \ldots du_{m+1} =$$

$$= \sum_{i=1}^{m+1} \int_{K_i} [a_i(u_1, \ldots, \beta_i, \ldots, u_{m+1}) - (a_i(u_1, \ldots, \alpha_i, \ldots, u_{m+1})] \cdot$$

$$\cdot\, du_1 \ldots \widehat{du_i} \ldots du_{m+1} = 0.$$

Corolário. *A integral de uma forma exata sobre uma superfície orientável compacta sem bordo é zero.*

Segue-se daí que se α é uma forma de grau m e classe C^1 numa superfície N (de dimensão qualquer, com ou sem bordo, orientável ou não) e M é uma superfície compacta orientável, sem bordo, contida em N, então $\int_M d\alpha = 0$. Freqüentemente, N é um aberto do espaço euclidiano. Como sabemos, $\int_M d\alpha$ significa, mais precisamente, $\int_M i^*(d\alpha)$, onde $i \colon M \to N$ é a inclusão.

As superfícies compactas sem bordo são chamadas *superfícies fechadas*. Assim, o corolário acima estende o Teorema 9 do Capítulo IV, segundo o qual a integral de uma forma exata de grau 1 ao longo de um caminho fechado é zero. Naquele teorema também está incluída uma recíproca, a qual está contida no Exercício 8.20, para $N = \mathbb{R}^m - \{0\}$.

Resulta ainda do corolário acima que se ω é uma forma positiva contínua, de grau m, numa superfície orientada compacta m-dimensional sem bordo M (ou, mais geralmente, se $\omega \geq 0$ mas não é $\omega \equiv 0$) então ω é uma forma fechada que não é exata em M. Exemplo: o elemento de volume.

Se M não é compacta (ou se é compacta mas possui bordo), então a forma contínua positiva de grau máximo em M pode ser exata. Por exemplo, se $M = \mathbb{R}^m$ (ou M é uma bola fechada em $\mathbb{R}^m$) então a forma

$\omega = dx_1 \wedge \cdots \wedge dx_m$ é positiva, mas sabemos que $\omega = d\alpha$, onde

$$\alpha = \frac{1}{m} \sum_{i=1}^{m} (-1)^{i+1} x_i dx_1 \wedge \cdots \wedge \widehat{dx_i} \wedge \cdots \wedge dx_m .$$

Teorema de Stokes. *Seja ω uma forma diferencial de classe C^1, de grau m, com suporte compacto, numa superfície orientada M, de dimensão $m+1$, cujo bordo ∂M é munido da orientação induzida. Então $\int_M d\omega = \int_{\partial M} \omega$.*

Demonstração: No enunciado acima, $\int_{\partial M} \omega$ significa a integral da restrição de ω ao bordo ∂M, isto é, da forma $i^*\omega$, onde $i \colon \partial M \to M$ é a aplicação de inclusão. Se $\omega = \omega_1 + \cdots + \omega_k$ e o teorema vale para cada parcela ω_i então vale para a soma ω, pois $d\omega = \Sigma\, d\omega_i$ e

$$\int_M d\omega = \sum_{i=1}^{k} \int_M d\omega_i = \sum_{i=1}^{k} \int_{\partial M} \omega_i = \int_{\partial M} \omega.$$

Ora, usando uma partição da unidade, podemos escrever ω como soma finita de formas que têm suportes compactos, contidos em vizinhanças parametrizadas. Logo, não há perda de generalidade em admitir que o suporte de ω está contido na imagem da parametrização positiva $\varphi \colon U_0 \to U$. Se $U \cap \partial M = \varnothing$ então $\int_{\partial M} \omega = 0$ e, pelo teorema anterior, $\int_M d\omega = 0$. Basta então considerar o caso em que $\partial U = U \cap \partial M \neq \varnothing$. Suporemos inicialmente que $m > 0$, ou seja, que dim. $M \geq 2$. Então a parametrização φ pode ser admitida padronizada, isto é, definida no aberto U_0 do semi-espaço $H_0 = \{(u_0, u_1, \ldots, u_m) \in \mathbb{R}^{m+1}; u_0 \leq 0\}$. Pela definição de orientação induzida, a restrição de φ a $\partial U_0 = U_0 \cap \partial H_0$ será uma parametrização positiva em ∂M, cuja imagem é ∂U. Para cada $x = \varphi(u) \in U$ escrevamos

$$\omega(x) = \sum_{i=0}^{m} (-1)^i\, a_i(u) du_0 \wedge \cdots \wedge \widehat{du_i} \wedge \cdots \wedge du_m ,$$

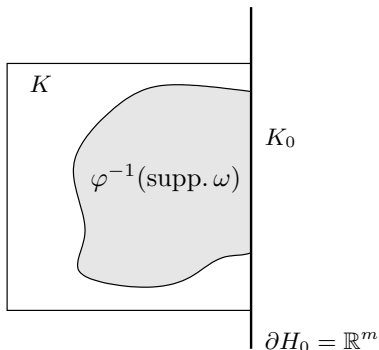
donde

$$(d\omega)(x) = \left[\sum_{i=0}^{m} \frac{\partial a_i}{\partial u_i}(u) \right] du_0 \wedge \cdots \wedge du_m .$$

Tomemos um bloco $K = \prod_{i=0}^{m} [\alpha_i, \beta_i]$ contendo $\varphi^{-1}(\text{supp.}\, \omega)$, com $\beta_0 = 0$, o que garante $K \subset H_0$. Para cada $i = 0, 1, \ldots, m$, seja K_i o produto cartesiano dos intervalos $[\alpha_j, \beta_j]$, com $j \neq i$. Em particular, $K_0 = \prod_{i=1}^{m} [\alpha_i, \beta_i]$

é um bloco em $\mathbb{R}^m \quad (= \partial H_0)$ contendo $\varphi^{-1}(\text{supp.}\, i^*\omega)$. Consideramos as funções $a_0, \ldots, a_m$ definidas (e contínuas) em K, pelo processo habitual de atribuir-lhes o valor zero nos pontos de $K - U_0$. Para todo $x = \varphi(u) = \varphi(0, u_1, \ldots, u_m) \in \partial U$, é evidente que $(i^*\omega)(x) = a_0(0, u_1, \ldots, u_m) du_1 \wedge \cdots \wedge du_m$, logo

$$\int_{\partial M} \omega = \int_{\partial M} i^*\omega = \int_{K_0} a_0(0, u_1, \ldots, u_m) du_1 \ldots du_m \,.$$

Se, para algum $i > 0$, a i-ésima coordenada do ponto $u \in K$ é igual a α_i ou β_i, todas as funções a_j se anulam no ponto u pois $\varphi(u)$ não pertence ao interior de supp. ω. Isto também ocorre se $u_0 = \alpha_0$, mas, evidentemente, nada se pode dizer sobre $a_0(0, u_1, \ldots, u_m)$. Reduzindo uma integral múltipla a integrais repetidas, podemos escrever

$$\int_M d\omega = \sum_{i=0}^{m} \int_K \frac{\partial a_i}{\partial u_i}(u) du_0 du_1 \ldots du_m =$$

$$= \sum_{i=0}^{m} \int_{K_i} \left[\int_{\alpha_i}^{\beta_i} \frac{\partial a_i}{\partial u_i}(u) du_i \right] du_0 \ldots \widehat{du_i} \ldots du_m =$$

$$= \sum_{i=0}^{m} \int_{K_i} [a_i(u_0, \ldots, \beta_i, \ldots, u_m) - a_i(u_0, \ldots, \alpha_i, \ldots, u_m)] \cdot$$

$$\cdot\, du_0 \ldots \widehat{du_i} \ldots du_m = \int_{K_0} a_0(0, u_1, \ldots, u_m) du_1 \ldots du_m = \int_{\partial M} \omega.$$

Resta considerar o caso em que $m = 0$, ou seja, em que M é uma superfície orientada de dimensão 1 e ω, tendo grau zero, se reduz a uma

função $f\colon M \to \mathbb{R}$, de classe C^1. Devemos provar que $\int_M df = \int_{\partial M} f$. Mas o que significa a integral de uma função f (com suporte compacto) sobre uma "superfície" ∂M de dimensão zero? Se ∂M é o conjunto dos pontos isolados $x_1, x_2, \ldots$ então $\int_{\partial M} f = \sum\limits_i \int_{\{x_i\}} f$ (soma finita porque f tem suporte compacto) e $\int_{\{x_i\}} f = \pm f(x_i)$, por definição, onde se toma o sinal $+$ ou o sinal $-$ conforme o sinal atribuido a x_i na orientação induzida por M. Mais explicitamente, se o vetor de uma base positiva de M no ponto x_i aponta para fora de M então $\int_{\{x_i\}} f = f(x_i)$. Se, ao contrário, a base positiva $T_{x_i}M$ aponta para dentro de M, então $\int_{\{x_i\}} f = -f(x_i)$.

Uma vez dada a definição acima, a demonstração anterior se adapta imediatamente ao caso dim. $M = 1$. Se as bases positivas de T_xM apontam para fora de M no ponto $x \in \partial M$, tomamos uma parametrização positiva $\varphi\colon (\alpha, 0] \to U$ de uma vizinhança U de x em M e então, se supp. f é um compacto contido em U, temos $f(\varphi(\alpha)) = 0$, logo

$$\int_M df = \int_\alpha^0 f(\varphi(t))\varphi'(t)dt = f(\varphi(0)) - f(\varphi(\alpha)) =$$
$$= f(\varphi(0)) = f(x) = \int_{\partial M} f.$$

Se, porém, as bases positivas de T_xM apontam para dentro de M no ponto x, as parametrizações positivas de vizinhanças de x com imagens contendo o suporte de f são da forma $\varphi\colon [0, \beta) \to U$, com $\varphi(0) = x$ e $f(\varphi(\beta)) = 0$. Então, como acima, temos

$$\int_M df = \int_0^\beta f'(\varphi(t))\varphi'(t)dt = f(\varphi(\beta)) - f(\varphi(0)) = -f(x) = \int_{\partial M} f.$$

É instrutivo reformular o Teorema de Stokes nos casos clássicos, que se referem à integral de um campo de vetores ao longo de uma hiperfície orientada.

Seja então $X = (a_0, a_1, \ldots, a_m)$, um campo de vetores de classe C^k, definido no aberto $U \subset \mathbb{R}^{m+1}$ por suas $m+1$ funções coordenadas $a_i\colon U \to \mathbb{R}$. Ao campo X associaremos a forma diferencial

$$\alpha_X = \sum_{i=0}^m (-1)^i\, a_i dx_0 \wedge \cdots \wedge \widehat{dx_i} \wedge \cdots \wedge dx_m\,,$$

de classe C^k no aberto U. O desenvolvimento de um determinante em relação à sua primeira coluna mostra que, para $x \in U$ e $w_1, \ldots, w_m \in \mathbb{R}^{m+1}$ arbitrários, vale $\alpha_X(x) \cdot (w_1, \ldots, w_m) = \det(X(x), w_1, \ldots, w_m)$, onde à direita temos o determinante da matriz $(m+1) \times (m+1)$ cujas colunas são os vetores $X(x)$, $w_1, \ldots, w_m$.

Se $M \subset U$ é uma hiperfície orientada e o campo X tem suporte compacto, definimos no parágrafo 6 a integral $\int_M X$ como $\int_M \langle X, v\rangle\omega$, onde ω é a forma elemento de volume de M e v é o campo de vetores unitários normais que define a orientação de M. [Isto é, para cada $x \in M$, $v(x)$ tem comprimento 1, é perpendicular a T_xM e, se $\{w_1, \ldots, w_m\} \subset T_xM$ é uma base positiva, tem-se $\det. (v(x), w_1, \ldots, w_m) > 0$.] Esta definição reduz a integral $\int_W X$ à integral de uma forma diferencial, a saber, a forma $x \mapsto \langle X(x), v(x)\rangle \cdot \omega(x)$.

Inicialmente, observamos que, em cada ponto $x \in M$, esta forma é igual a $\alpha_X(x)$, conforme definida acima. De fato, dada qualquer base positiva $\{w_1, \ldots, w_m\} \subset T_xM$, temos

$$\begin{aligned} \alpha_X(x) \cdot (w_1, \ldots, w_m) &= \det(X(x), w_1, \ldots, w_m) = \\ &= \langle X(x), w_1 \times \cdots \times w_m\rangle = \\ &= \langle X(x), v(x)\rangle \cdot |w_1 \times \cdots \times w_m| = \\ &= \langle X(x), v(x)\rangle \cdot \omega(x) \cdot (w_1, \ldots, w_m), \end{aligned}$$

pois o produto vetorial $w_1 \times \cdots \times w_m$ é um vetor normal a M, na direção da normal positiva $v(x)$. Segue-se que $\alpha_X = \langle X, v\rangle \cdot \omega$ ao longo de M.

A notação tradicional para a forma elemento de volume numa superfície M é dM.

Podemos então dizer que se $X = (a_0, \ldots, a_m)$ é um campo de vetores com suporte compacto, contínuo no aberto $U \subset \mathbb{R}^{m+1}$, e $M \subset U$ é uma hiperfície orientada, então

$$\int_M \langle X, v\rangle \cdot dM = \int_M \sum_{i=0}^{m} (-1)^i \, a_i dx_0 \wedge \cdots \wedge \widehat{dx_i} \wedge \cdots \wedge dx_m \,.$$

Há um caso particular importante, em que X é um campo de classe C^1 no aberto $U \subset \mathbb{R}^{m+1}$ e $\Omega \subset U$ é o que se costuma chamar um *domínio compacto com fronteira regular* de classe C^k $(k \geq 1)$. Isto significa que Ω é uma superfície com bordo, compacta, de dimensão $m+1$ e classe C^k, contida em U. [Neste caso, a orientação de $\mathbb{R}^{m+1}$ induz naturalmente uma orientação em Ω pois, para cada $x \in \Omega$, temos

$T_x\Omega = \mathbb{R}^{m+1}$.] O interior de Ω é um subconjunto aberto limitado em $\mathbb{R}^{m+1}$ e o bordo $\partial\Omega$ (que é também a fronteira de int. Ω em $\mathbb{R}^{m+1}$) é uma hiperfície compacta, orientada, em $\mathbb{R}^{m+1}$. A diferencial da forma α_X é $d\alpha_X = \left(\sum_{i=0}^{m} \frac{\partial a_i}{\partial x_i}\right) dx_0 \wedge \cdots \wedge dx_m$. A função div. $X \colon U \to \mathbb{R}$, definida por

$$(\text{div.}\, X)(x) = \frac{\partial a_0}{\partial x_0}(x) + \cdots + \frac{\partial a_m}{\partial x_m}(x),$$

chama-se a *divergência* do campo X. O Teorema de Stokes nos permite, portanto, afirmar que, nestas condições, se $M = \partial\Omega$, então

$$\int_M \langle X, v\rangle \cdot dM = \int_\Omega \text{div.}\, X \cdot dx,$$

onde, no segundo membro, temos a integral ordinária da função contínua div. X sobre o compacto J-mensurável $\Omega \subset \mathbb{R}^{m+1}$.

A fórmula acima é conhecida como o *Teorema da Divergência*, de Gauss e Ostrogradski.

Seja agora $X = (a, b, c)$ um campo de classe C^1 no aberto $U \subset \mathbb{R}^3$, que contém a superfície compacta orientada M (de dimensão 2), cujo bordo C munimos da orientação induzida. Ao campo X associamos a forma $\beta_X = adx + bdy + cdz$, de grau 1 e classe C^1 em U. O Teorema de Stokes $\int_M d\beta_X = \int_C \beta_X$ se escreve, explicitamente:

$$\int_M \left(\frac{\partial c}{\partial y} - \frac{\partial b}{\partial z}\right) dy \wedge dz + \left(\frac{\partial a}{\partial z} - \frac{\partial c}{\partial x}\right) dz \wedge dx + \\ + \left(\frac{\partial b}{\partial x} - \frac{\partial a}{\partial y}\right) dx \wedge dy = \int_C adx + bdy + cdz.$$

Para exprimir a igualdade acima em termos de vetores (em vez de formas) introduzimos o *rotacional* do campo X, que é o novo campo rot: $U \to \mathbb{R}^3$, dado por

$$\text{rot.}X = \left(\frac{\partial c}{\partial y} - \frac{\partial b}{\partial z}, \frac{\partial a}{\partial z} - \frac{\partial c}{\partial x}, \frac{\partial b}{\partial x} - \frac{\partial a}{\partial y}\right).$$

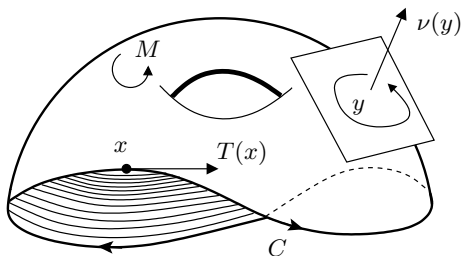
O primeiro membro da igualdade se escreve então $\int_M \langle \text{rot.}X, v\rangle \cdot dM$, onde v é o campo de vetores unitários normais a M, definidos pela orientação de M, e dM é o elemento de área.

Quanto ao segundo membro, seja ds a forma elemento de arco em C. Para $x \in C$ e $v \in T_xC$, temos $ds \cdot v = \pm|v|$, conforme v aponte no

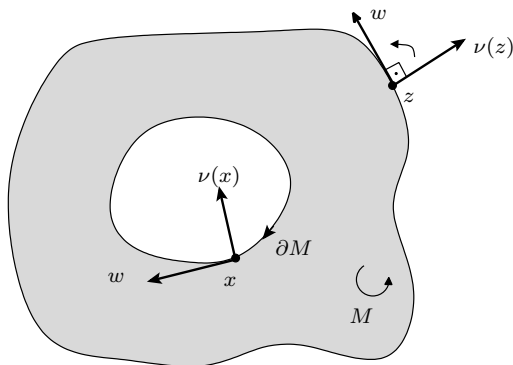
sentido positivo de C ou não. Analogamente, se $\tau(x) \in T_xC$ representa o vetor unitário tangente que aponta no sentido positivo de C, teremos $\tau(x) = \pm v/|v|$ (conforme a orientação de v), para todo $v \neq 0$ em T_xC. O segundo membro da igualdade em questão é igual a $\int_C \beta_X$. Ora,

$$0 \neq v \in T_xC \Rightarrow \beta_X \cdot v = \langle X, v\rangle = \left\langle X, \frac{v}{|v|}\right\rangle \cdot |v| =$$
$$= \langle X, \pm\tau\rangle \cdot \pm ds \cdot v = \langle X, \tau\rangle(ds \cdot v).$$

Assim, $\beta_X = \langle X, \tau\rangle \cdot ds$ e podemos escrever

$$\int_M \langle \text{rot.}X, v\rangle \cdot dM = \int_C \langle X, \tau\rangle ds$$

Este é o *Teorema de Stokes* da Análise Vetorial. O primeiro membro representa o *fluxo* do campo rot.X através da superfície M e o segundo membro é a *circulação* do campo X ao longo do bordo $C = \partial M$.

A importância das fórmulas acima, expressas em linguagem de campos de vetores, resulta das interpretações físicas que podem ser dadas ao fluxo, à circulação, etc. O leitor interessado pode consultar Courant [2].

A mais simples das três fórmulas integrais da Análise Vetorial clássica é o Teorema de Green (às vezes chamado de Teorema de Riemann-Gauss). Ele se refere a uma superfície com bordo, compacta de classe C^1 e dimensão 2 em $\mathbb{R}^2$, isto é, a um domínio compacto $M \subset \mathbb{R}^2$ com fronteira regular de classe C^1. O domínio M tem a orientação natural de $\mathbb{R}^2$ e seu bordo ∂M a orientação induzida: em cada ponto $x \in \partial M$, um vetor tangente não nulo $w \in T_x(\partial M)$ aponta na direção positiva se, e somente se, $\{v(x), w\}$ é uma base positiva de $\mathbb{R}^2$, onde $v(x)$ é a normal unitária que aponta para fora de M.

Sejam $f, g \colon M \to \mathbb{R}$ funções de classe C^1. O *Teorema de Green* diz que

$$\int_M \left(\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y} \right) dxdy = \int_{\partial M} f dx + g dy.$$

Isto é simplesmente o Teorema de Stokes aplicado à forma diferencial $\beta = f dx + g dy$, definida na superfície com bordo M. O primeiro membro é uma integral dupla sobre o compacto J-mensurável M e o segundo membro é uma integral curvilínea.

O Teorema de Green pode ser considerado como um caso particular do Teorema da Divergência [tome o campo $X = (g, -f)$] ou do teorema do rotacional [tome $X = (f, g, 0)$]. Também o Teorema Fundamental do Cálculo, $\int_b^a f'(x)dx = f(b) - f(a)$ é um caso particular do Teorema de Stokes.

Exemplo 26. Como primeira aplicação do Teorema de Stokes, vamos mostrar que não pode existir uma retração de classe C^2 de uma superfície orientada, compacta, M sobre seu bordo ∂M. Em particular, a esfera S^n não é um retrato de classe C^∞ da bola B^{n+1}, o que nos dá uma nova demonstração do argumento crucial para estabelecer o Teorema do Ponto Fixo de Brouwer. (Veja §5.) [Na realidade, pode-se provar que não existe sequer uma retração contínua de uma superfície compacta sobre seu bordo. Vamos supor C^2, para que a forma induzida seja C^1.] Com efeito, admitamos, por absurdo, a existência de uma retração $f \colon M \to \partial M$, de classe C^2. Seja ω uma forma C^1 de grau máximo em ∂M, tal que $\int_{\partial M} \omega \neq 0$. (Por exemplo, ω pode ser o elemento de volume de ∂M.) Por ter grau igual à dimensão de ∂M, ω é fechada, isto é, $d\omega = 0$. Pelo Teorema de Stokes, mais o fato de que $f|\partial M =$ identidade, temos

sucessivamente:

$$0 \neq \int_{\partial W} \omega = \int_{\partial W} f^* \omega = \int_M d(f^* \omega) = \int_M f^*(d\omega) = \int_M f^* 0 = 0,$$

um absurdo.

A próxima aplicação do Teorema de Stokes é o teorema abaixo, que estende para dimensão > 1 o Teorema 12 do Capítulo IV, segundo o qual a integral de uma 1-forma fechada sobre um caminho fechado é invariante por homotopias desse caminho.

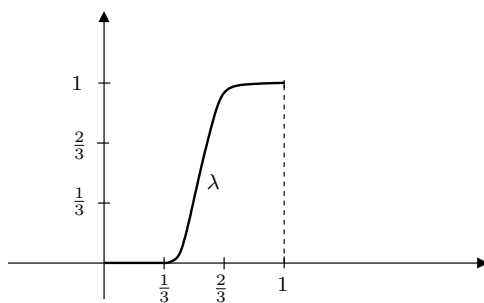
Uma *homotopia* entre as aplicações $f, g \colon M \to N$ (de classe C^k) é uma aplicação contínua $H \colon I \times M \to N$ onde $I = [0,1]$ e, para todo $x \in M$, tem-se $H(0,x) = f(x)$, $H(1,x) = g(x)$. Escreve-se $H \colon f \simeq g$ ou, simplesmente $f \simeq g$, para indicar que existe uma homotopia H entre f e g. Se $H \in C^k$, diz-se que f e g são C^k-*homotópicas*.

Por exemplo, se $B \subset \mathbb{R}^n$ é a bola fechada unitária, então $H \colon I \times B \to B$, $H(t,x) = (1-t)x$, é uma homotopia C^∞ entre a aplicação identidade de B e a aplicação constante $x \mapsto 0$. Diz-se que um conjunto $X \subset \mathbb{R}^n$ é *contrátil* quando existe uma homotopia $H \colon I \times X \to X$ entre a aplicação identidade de X e uma aplicação constante.

É evidente que a relação "f e g são C^k-homotópicas"é reflexiva e simétrica. Se $k = 0$, é também imediato que ela é transitiva pois se $H \colon f \simeq g$ e $K \colon g \simeq h$ são homotopias de classe C^0, obtemos uma homotopia $L \colon I \times M \to N$, de classe C^0 entre f e h, pondo, para todo $t \in I$ e todo $x \in M$, $L(t,x) = H(2t,x)$ se $0 \leq t \leq 1/2$ e $L(t,x) = K(2t-1,x)$ quando $1/2 \leq t \leq 1$. Se H e K são homotopias de classe C^1, a homotopia L, assim definida, pode não ser de classe C^1, mesmo quando a superfície M se reduz a um ponto. [A justaposição de dois caminhos C^1 é, em geral, apenas um caminho C^0.]

A transitividade da relação de homotopia em classe C^k $(k \geq 1)$ torna-se fácil de provar (e a demonstração acima se aplica ipsis literis) quando supomos que as homotopias $H \colon f \simeq g$ e $K \colon g \simeq h$ (de classe C^k) são adaptadas. Uma homotopia $H \colon I \times M \to N$ diz-se *adaptada* quando existe ε $(0 < \varepsilon < 1/2)$ tal que, para todo $x \in M$, $H(t,x)$ é constante nos intervalos $0 \leq t < \varepsilon$ e $1 - \varepsilon < t \leq 1$. Noutras palavras: $H(t,x) = f(x)$ para todo $t \in [0,\varepsilon)$ e $H(t,x) = g(x)$ seja qual for $t \in (1-\varepsilon, 1]$, onde $x \in M$ e qualquer. Quando $H \colon f \simeq g$ e $K \colon g \simeq h$ são homotopias adaptadas de classe C^k então a homotopia $L \colon I \times M \to N$, definida por $L(t,x) = H(2t,x)$ para $t \in [0,1/2]$ e $L(t,x) = K(2t-1,x)$ se $t \in [1/2,1]$, é de classe C^k, como se vê facilmente.

Resta mostrar que se existe uma homotopia $H\colon I\times M\to N$ de classe C^k entre f e g, existe também uma homotopia adaptada $K\colon f\simeq g$. Para isto, basta tomar uma função $\lambda\colon [0,1]\to[0,1]$, de classe C^∞, tal que $\lambda(t)=0$ se $t\in[0,1/3]$, $\lambda(t)=1$ se $t\in[2/3,1]$ e pôr $K(t,x)=H(\lambda(t),x)$ para todo $t\in I$ e todo $x\in M$. [Para obter λ podemos usar, por exemplo, a função δ, definida no início da seção 4, pondo $\lambda(t)=\delta(3t-3)$.]

Concluimos portanto que a relação "existe uma homotopia de classe C^k entre f e g"é uma equivalência no conjunto das aplicações de classe C^k de uma superfície M numa superfície N.

Feitas essas considerações preliminares, vamos ao

Teorema 17. *Sejam M uma superfície de dimensão m, compacta, orientada, sem bordo, e $f,g\colon M\to N$ aplicações C^2-homotópicas. Se ω é uma forma de classe C^1, fechada, de grau m, em N então $\int_M f^*\omega = \int_M g^*\omega$.*

Demonstração: Seja $H\colon I\times M\to N$ uma homotopia de classe C^2 entre f e g. Sabemos que $\partial(I\times M)=M_1\cup(-M_0)$, onde $M_1=\{1\}\times M$ e $M_0=\{0\}\times M$. Como $H(0,x)=f(x)$ e $H(1,x)=g(x)$ para todo $x\in M$, podemos escrever (usando o Teorema de Stokes)

$$\int_M g^*\omega-\int_M f^*\omega=\int_{M_1}H^*\omega-\int_{M_0}H^*\omega=\int_{M_1\cup(-M_0)}H^*\omega=$$
$$=\int_{\partial(I\times M)}H^*\omega=\int_{I\times M}d(H^*\omega)=\int_{I\times M}H^*(d\omega)=0,\ \text{pois } d\omega=0.$$

Corolário. *Uma superfície orientável M, compacta, sem bordo, não é C^2-contrátil.*

De fato, se existisse uma homotopia $H\colon I\times M\to M$, de classe C^2, entre a aplicação identidade e uma aplicação constante, então, pelo

teorema acima, teríamos $\int_M \omega = 0$ para toda forma de grau máximo em M, o que não é verdade.

Observação: Toda aplicação contínua $f\colon M \to N$, entre superfícies de classe C^k, é homotópica a uma aplicação $g\colon M \to N$, de classe C^k. Isto é trivial no caso em que $N = \mathbb{R}^n$ pois $H(t,x) = (1-t)f(x)$ é, então, uma homotopia entre f e a aplicação constante zero. Para uma aplicação contínua $f\colon M \to S^n$ a afirmação também pode ser facilmente provada assim: aproximamos f por uma aplicação $g\colon M \to S^n$, de classe C^k, tal que $|f(x) - g(x)| < 2$ para todo $x \in M$. (Cfr. §5.) Então, para todo $x \in M$, os pontos $f(x), g(x) \in S^n$ não são antípodas, ou seja, o segmento de reta $[f(x), g(x)]$ não contém 0. Portanto, a aplicação $H\colon I \times M \to S^n$, dada por

$$(*) \qquad H(t,x) = \frac{(1-t)f(x) + tg(x)}{|(1-t)f(x) + tg(x)|} \ .$$

é contínua, logo é uma homotopia entre f e g. No caso geral (que não provaremos) faz-se uso de uma "vizinhança tubular"de N no espaço euclidiano que a contém. (Veja [13].)

Além disso, se $f, g\colon M \to N$ são de classe C^k e existe uma homotopia contínua $H\colon f \simeq g$, existe também uma homotopia de classe C^k entre f e g. Novamente, provaremos este fato no caso em que N é a esfera S^n. Aproximamos a homotopia $H\colon 1 \times M \to S^n$ por uma aplicação $K\colon I \times M \to S^n$, de classe C^k, tal que $|K(t,x) - H(t,x)| < 2$ para todo $t \in I$ e todo $x \in M$, K é uma homotopia de classe C^k entre duas aplicações $f_1, g_1\colon M \to S^n$, de classe C^k, com $|f(x) - f_1(x)| < 2$ e $|g(x) - g_1(x)| < 2$ para todo $x \in M$. Como vimos acima, isto nos permite concluir que $f \simeq f_1$ e $g \simeq g_1$. A fórmula (*) acima exibida mostra claramente que estas duas homotopias são de classe C^k. Pela transitividade da homotopia C^k, concluimos que existe uma homotopia de classe C^k entre f e g.

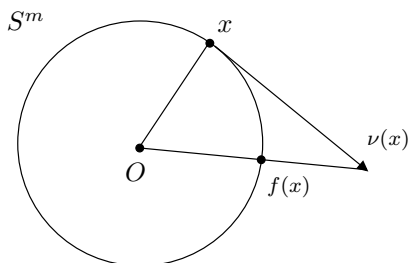
Das observações acima resulta ainda que se uma superfície de classe C^k é contrátil em classe C^0 então ela é também contrátil em classe C^k.

Estas considerações têm a finalidade de mostrar que o teorema de aproximação de aplicações contínuas por aplicações diferenciáveis pode ser utilizado eficazmente para reduzir questões topológicas (referentes a aplicações contínuas) a questões diferenciáveis, que podem ser tratadas pelos instrumentos da Análise. Este é o método da Topologia Diferencial.

Exemplo 27. Como aplicação do Teorema 17, vamos mostrar que a aplicação antípoda $\alpha\colon S^m \to S^m$, $\alpha(x) = -x$, é homotópica à aplicação identidade de S^m se, e somente se, m é ímpar. Em primeiro lugar, se $m = 2k - 1$ é ímpar, os pontos de $S^m = S^{2k-1} \subset \mathbb{R}^{2k}$ têm um número par de coordenadas, logo podem ser escritos sob a forma $z = (z_1, z_2, \ldots, z_k)$, onde cada z_i é um número complexo, com $\sum_{i=1}^{k} |z_i|^2 = 1$. Obtemos então uma homotopia $H\colon I \times S^m \to S^m$, entre a identidade e α, pondo $H(t, z) = e^{t\pi i} \cdot z = (e^{t\pi i} \cdot z_1, \ldots, e^{t\pi i} \cdot z_k)$. Reciprocamente, se existir uma homotopia $H\colon I \times S^m \to S^m$ entre a aplicação identidade e a aplicação antípoda α, pela Observação acima, podemos supor que H é de classe C^∞. Consideremos qualquer forma contínua ω em S^m tal que $\int_{S^m} \omega \neq 0$ (por exemplo, a forma elemento de volume). Pelo Exemplo 41, Capítulo V, sabemos que o difeomorfismo $\alpha\colon S^m \to S^m$ preserva orientação se, e somente se, m é ímpar. Portanto, $\int_{S^m} \alpha^*\omega = (-1)^{m+1} \int_{S^m} \omega$. Por outro lado, a homotopia $\alpha \simeq \mathrm{id}_{S^m}$ nos dá, pelo Teorema 17, $\int_{S^m} \alpha^*\omega = \int_{S^m} \omega$. Segue-se que $(-1)^{m+1} = 1$, ou seja, m é ímpar.

Lembremos inicialmente que se $M \subset \mathbb{R}^n$ é uma superfície e $v\colon M \to \mathbb{R}^n$ é um campo de vetores tangentes (isto é, $v(x) \in T_xM$ para todo $x \in M$), um ponto $x \in M$ chama-se uma *singularidade* de v quando $v(x) = 0$. Temos então o

Teorema 18. (Poincaré e Brouwer). *Se m é par, todo campo contínuo de vetores tangentes à esfera S^m possui ao menos uma singularidade.*

Demonstração: Suponhamos que $v\colon S^m \to \mathbb{R}^{m+1}$ seja um campo contínuo de vetores tangentes à esfera S^m, com $v(x) \neq 0$ para todo $x \in S^m$. A partir do campo v, definimos a aplicação contínua $f\colon S^m \to S^m$, pondo $f(x) = [x + v(x)]/|x + v(x)|$, para todo $x \in S^m$. Como $v(x) \neq 0$, temos $f(x) \neq x$ para todo x, isto é, f não possui pontos fixos.

Isto nos permite definir uma homotopia $H\colon I \times S^m \to S^m$, entre f e a aplicação antípoda de S^m, pondo

$$H(t,x) = \frac{(1-t)f(x) - tx}{|(1-t)f(x) - tx|}\,.$$

[O denominador acima nunca se anula porque isto significaria que 0 pertence ao segmento de reta $[f(x), -x]$ e, como $|f(x)| = 1$, teríamos $f(x) = x$.] Por outro lado, a aplicação $K\colon I \times S^m \to S^m$, dada por $K(t,x) = [x + tv(x)]/|x + tv(x)|$, é uma homotopia entre f e a aplicação identidade de S^m. Por transitividade, concluimos que a aplicação antípoda $\alpha\colon S^m \to S^m$ é homotópica à identidade, logo m é ímpar.

O Teorema 17, como está enunciado, não vale no caso em que a superfície M não é compacta. Por exemplo, o espaço euclidiano é uma superfície orientável, sem bordo, contrátil em classe C^∞. Dito de outro modo: a aplicação identidade e a aplicação constante (igual a 0) são homotópicas em classe C^∞ mas não se tem $\int_{\mathbb{R}^m} \omega = 0$ para toda m-forma ω com suporte compacto.

Para estender o Teorema 17, introduziremos uma definição que tem provado ser útil quando se tem uma aplicação cujo domínio não é compacto.

Sejam X, Y subconjuntos de espaços euclidianos. Uma aplicação contínua $f\colon X \to Y$ diz-se *própria* quando a imagem inversa $f^{-1}(K)$ de todo compacto $K \subset Y$ é um conjunto compacto.

Por exemplo, se X é compacto, toda aplicação contínua $f\colon X \to Y$ é própria. Já se Y é compacto, $f\colon X \to Y$ só pode ser própria quando $X = f^{-1}(Y)$ for compacto. Um homeomorfismo $f\colon X \to Y$ é uma aplicação própria.

Evidentemente, uma *homotopia própria* é uma homotopia $H\colon I \times X \to Y$ que é também uma aplicação própria. Daí resulta, em particular, que cada aplicação $H_t\colon X \to Y$, $H_t(x) = H(t,x)$, $t \in I$, é própria, pois se $K \subset Y$ é compacto então $H_t^{-1}(K) = H^{-1}(K) \cap [\{t\} \times X]$ é um subconjunto fechado do compacto $H^{-1}(K)$. Portanto, se H é uma homotopia própria entre f e g então f e g são aplicações próprias. Diremos que f e g são *propriamente homotópicas*.

A reformulação do Teorema 17 pode ser agora enunciada:

Teorema 17a. *Sejam $f, g\colon M \to N$ aplicações propriamente homotópicas em classe C^2, definidas na superfície orientada, sem bordo, M, de*

dimensão m. Se ω é uma forma fechada de grau m, classe C^1 e suporte compacto em N então $\int_M f^*\omega = \int_M g^*\omega$.

Demonstração: Se $H\colon I \times M \to N$ é uma homotopia própria de classe C^2 entre f e g então $H^*\omega$, $f^*\omega$ e $g^*\omega$ têm suportes compactos em M, o mesmo ocorrendo com $d\omega$ e $H^*(d\omega)$. [Por exemplo, supp. $H^*\omega$ é um subconjunto fechado de H^{-1} (supp.ω).] Com esta observação, a demonstração do Teorema 17 se aplica tal qual foi feita acima.

Os lemas abaixo visam dar alguma familiaridade com o conceito de aplicação própria.

Lema 1. *Toda aplicação própria $f\colon X \to Y$ é fechada.*

Demonstração: Seja $F \subset X$ fechado em X. Devemos provar que $f(F)$ é fechado em Y. Com efeito se uma seqüência de pontos $f(x_k) \in f(F)$ converge para um ponto $y \in Y$, o conjunto K, formado pelos $f(x_k)$ e mais o ponto y, é compacto, logo a seqüência de pontos $x_k \in F$ está contida no compacto $f^{-1}(K) \subset X$, portanto possui uma subseqüência (x'_k) que converge para um ponto $x \in f^{-1}(K) \subset X$. Como F é fechado em X, temos $x \in F$, portanto $y = \lim f(x'_k) = f(x) \in f(F)$. Assim, $f(F)$ é fechado em Y.

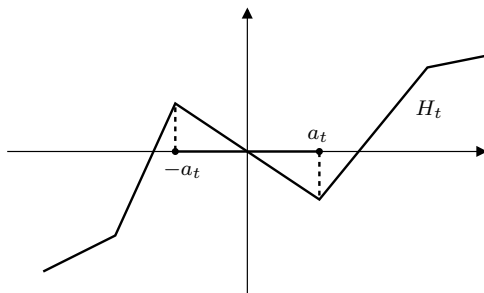
Para enunciar o lema seguinte, precisamos de uma definição. Dado $X \subset \mathbb{R}^m$, dizemos que uma seqüência de pontos $x_k \in Y$ *tende para o infinito em* X, e escrevemos $\lim x_k = \infty$ em X, quando (x_k) não possui subseqüências que convirjam para pontos de X. Se $X = \mathbb{R}^n$, "$\lim x_k = \infty$ em X "significa "$\lim |x_k| = +\infty$".

Lema 2. *As seguintes afirmações a respeito de uma aplicação contínua $f\colon X \to Y$ são equivalentes:*

1) *Para todo compacto $K \subset Y$, $f^{-1}(K)$ é compacto;*
2) *Se* $\lim x_k = \infty$ *em X então* $\lim f(x_k) = \infty$ *em Y.*

Demonstração: 1) $\Rightarrow$ 2). Com efeito, se alguma subseqüência $f(x'_k)$ convergisse para um ponto $y \in Y$ então o conjunto K, formado pelos pontos $f(x'_k)$ e mais y, seria compacto. Então a seqüência (x'_k) estaria contida no compacto $f^{-1}(K) \subset X$, logo possuiria uma subseqüência convergente em X. Para provar que 2) $\Rightarrow$ 1), seja $K \subset Y$ compacto. Toda seqüência de pontos $x_k \in f^{-1}(K)$ possui uma subsequência (x'_k) que converge para um ponto $x \in X$, já que sua imagem $(f(x_k))$ está contida no compacto K, logo possui subseqüência convergente. Como $f(x) = \lim f(x'_k)$ e cada $f(x'_k) \in K$, temos $f(x) \in K$, donde $x \in f^{-1}(K)$. Assim, $f^{-1}(K)$ é compacto.

O Lema 2 é útil para visualizar aplicações próprias. Dele decorre, por exemplo, que todo polinômio não-constante (real ou complexo) é uma aplicação própria.

Exemplo 28. Duas aplicações contínuas quaisquer $f, g\colon X \to \mathbb{R}^n$ são sempre homotópicas mas, pelo Teorema 17a, um difeomorfismo $f\colon \mathbb{R}^m \to \mathbb{R}^m$, que inverta orientação [por exemplo, $f(x_1, \ldots, x_m) = (-x_1, x_2, \ldots, x_m)$] não é propriamente homotópico à aplicação identidade. Com efeito, seja, por exemplo: $\xi : \mathbb{R}^m \to \mathbb{R}$ a função auxiliar de classe C^∞ definida no início do parágrafo 4. Então $\omega = \xi \cdot dx_1 \wedge \cdots \wedge dx_m$ é uma forma de classe C^∞ e suporte compacto em $\mathbb{R}^m$, com $\int_{\mathbb{R}^m} \omega > 0$. Como f inverte orientação, temos $\int_{\mathbb{R}^m} f^*\omega < 0$. [Vide Observação que se segue ao Teorema 12.] Portanto não existe uma homotopia própria entre f e a identidade de $\mathbb{R}^m$. Entretanto, vale a pena observar que existe uma homotopia $H\colon I \times \mathbb{R} \to \mathbb{R}$, entre a aplicação identidade e a aplicação $x \mapsto -x$ tal que $H_t\colon \mathbb{R} \to \mathbb{R}$ é

própria, para cada $t \in [0,1]$. Isto mostra, em particular, que cada H_t pode ser própria sem que a homotopia H o seja, Esboçamos acima o gráfico de H_t $0 < t < 1$. Temos $H_0(x) = x$, $H_1(x) = -x$, para todo $x \in \mathbb{R}$. Se $0 < t < 1$, tomamos $a_t = t/(1-t)$ e, no intervalo $[-a_t, a_t]$, fazemos $H_t(x) = -x$. Fixamos $\varepsilon > 0$ e, fora do intervalo $[-a_t - \varepsilon, a_t + \varepsilon]$, fazemos $H_t(x) = x$. Nos pequenos intervalos $[-a_t - \varepsilon, -a_t]$ e $[a_t, a_t + \varepsilon]$ tomamos H_t linear de modo a torná-la contínua. Quando t varia de 0 a 1, H_t varia continuamente da identidade à função $x \mapsto -x$.

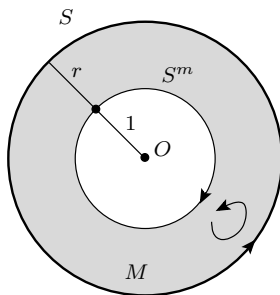
Exemplo 29. Seja $M \subset \mathbb{R}^2$ um domínio compacto com fronteira regular, isto é, uma superfície compacta de classe C^1 e dimensão 2, com bordo, no plano. Consideraremos em M a orientação natural, determinada por $\mathbb{R}^2$. A forma $\omega = \frac{1}{2}(xdy - ydx)$ e tal que $d\omega = dx \wedge dy$. Então Área de $M = \int_M dx \wedge dy = \int_M d\omega = \frac{1}{2} \int_{\partial M} xdy - ydx$, onde o bordo

∂M tem a orientação induzida por M. Esta fórmula tradicional, que exprime a área de uma região plana em termos de uma integral curvilínea ao longo do seu contorno, vale mais geralmente: se $M \subset \mathbb{R}^m$ é um domínio compacto m-dimensional com fronteira regular, então

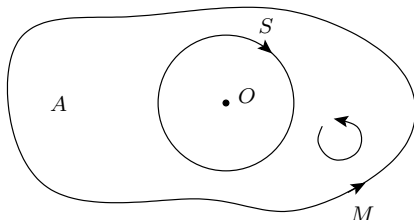
$$\text{vol. } M = \frac{1}{m}\int_{\partial M}\sum_{i=1}^{m}(-1)^{i+1}\, x_i dx_1 \wedge \cdots \wedge \widehat{dx_i} \wedge \cdots \wedge dx_m\,.$$

Exemplo 30. O elemento de ângulo sólido (vide Exemplo 17, §2) é uma forma α de classe C^∞ e grau m em $\mathbb{R}^{m+1} - \{0\}$ definida por

$$\alpha(x) = \frac{1}{|x|^{m+1}}\sum_{i=1}^{m+1}(-1)^{i+1}\, x_i dx_1 \wedge \cdots \wedge \widehat{dx_i} \wedge \cdots \wedge dx_{m+1}\,.$$

Sabemos que $d\alpha = 0$. (Vide final do §3.) Assim, a forma α é fechada. Mas α não é exata em $\mathbb{R}^{m+1} - \{0\}$. Com efeito, se existisse uma forma β em $\mathbb{R}^{m+1} - \{0\}$, de grau $m-1$ tal que $d\beta = \alpha$ então, para toda superfície compacta M, de dimensão m, sem bordo, contida em $\mathbb{R}^{m+1} - \{0\}$, teríamos $\int_M \alpha = \int_M d\beta = 0$, pelo Teorema 16. Mas $\int_{S^m} \alpha > 0$ pois, restrita à esfera unitária S^m, α coincide com a forma elemento de volume, logo é positiva, e sua integral também. Na realidade, se indicarmos com c_m o volume da esfera S^m [área (igual a 4π) se $m = 2$ e comprimento (igual a 2π) se $m = 1$], temos $\int_S \alpha = c_m$, seja qual for a esfera S, de raio arbitrário, com centro na origem $0 \in \mathbb{R}^{m+1}$. Com efeito, seja M a região compreendida entre S^m e S. (Por exemplo, se S tem raio $r > 1$ então $M = \{x \in \mathbb{R}^{m+1}; 1 \leq |x| \leq r\}$.) Então M é uma superfície com bordo. Considerando-se as orientações induzidas, temos $\partial M = S \cup (-S^m)$ se o raio de S é maior do que 1 e $\partial M = S^m \cup (-S)$ se o raio de S é menor do que 1. Então $0 = \int_M d\alpha = \int_{\partial M} \alpha = \int_S \alpha - \int_{S^m} \alpha$, o que prova a afirmação feita.

Mais geralmente, seja $M \subset \mathbb{R}^{m+1} - \{0\}$ uma hiperfície fechada (isto é, compacta, sem bordo). Existe um domínio compacto $A \subset \mathbb{R}^{m+1}$, com fronteira regular, tal que $\partial A = M$. (Ou seja, A é uma superfície compacta cujo bordo é M.) Este é o teorema (de Jordan-Brouwer) cuja demonstração não daremos aqui (v. [14] para o caso em que M é C^∞) mas na maioria das vezes em que M é descrita explicitamente, a existência de A é imediata. Admitindo-a, examinemos o valor da integral $\int_M \alpha$, onde α é o elemento de ângulo sólido em $\mathbb{R}^{m+1}$. Há duas possibilidades:

Primeira: $0 \notin A$, ou seja, a origem está "no exterior"de M. Então $\int_M \alpha = 0$. Com efeito, neste caso $A \subset \mathbb{R}^{m+1} - \{0\}$, logo a forma α está definida em todo o domínio A. Pelo Teorema de Stokes temos então $\int_M \alpha = \int_{\partial A} \alpha = \int_A d\alpha = 0$, pois $d\alpha = 0$.

Segunda: $0 \in A$, ou seja, a origem está "no interior"de M. Seja B uma bola de centro 0, disjunta de M, contida em A. O bordo de B é uma esfera S. Então $A - \text{int.}\, B = N$ é uma superfície

compacta, de dimensão $m+1$, com $\partial N = M \cup (-S)$, na qual a forma α está definida. Temos $0 = \int_N d\alpha = \int_{\partial N} \alpha = \int_M \alpha - \int_S \alpha = \int_M \alpha - c_m$. Segue-se que $\int_M \alpha = c_m$ (volume da esfera unitária de $\mathbb{R}^{m+1}$). Para $m = 3$, obtemos o *Teorema de Gauss*, segundo o qual, se $M \subset \mathbb{R}^3$ é uma superfície fechada então

$$\int_M \frac{xdy \wedge dz + ydz \wedge dx + zdx \wedge dy}{(x^2+y^2+z^2)^{3/2}} = 0 \text{ ou } 4\pi,$$

conforme a origem de $\mathbb{R}^3$ esteja fora ou dentro da superfície M.

Dado qualquer ponto $p = (a_1, \ldots, a_{m+1}) \in \mathbb{R}^{m+1}$, podemos considerar a forma elemento de ângulo sólido com vértice em p, definida por $\alpha_p = f_p^* \omega$ onde ω é o elemento de volume de S^m e $f_p \colon \mathbb{R}^{m+1} - \{0\} \to S^m$ é a projeção radial de vértice p, dada por $f_p(x) = (x-p)/|x-p|$. Tem-se,

para cada $x \in \mathbb{R}^{m+1} - \{p\}$,

$$\alpha_p(x) = \frac{1}{|x-p|^{m+1}} \sum_{i=1}^{m+1} (-1)^{i+1}(x_i - a_i) dx_1 \wedge \cdots \wedge \widehat{dx_i} \wedge \cdots \wedge dx_{m+1}\,.$$

Vale, analogamente, para toda hiperfície fechada $M \subset \mathbb{R}^{m+1} - \{p\}$, $\int_M \alpha_p = 0$ se p está fora do domínio compacto cujo bordo é M, enquanto que $\int_M \alpha_p = c_m$ (volume da esfera unitária S^m) se p pertence ao domínio compacto de $\mathbb{R}^{m+1}$ cujo bordo é M.

Exemplo 31. Conforme dissemos no parágrafo 6, o Teorema de Stokes não vale para formas com suportes não-compactos. Mostraremos isto por meio de um exemplo simples. O conjunto $M = \{(x,y) \in \mathbb{R}^2; 0 < x^2 + y^2 \leq 1\}$ é uma superfície não-compacta de dimensão 2, contida em $\mathbb{R}^2$, com $\partial M = S^1$. Consideremos em M a forma $\omega = (xdy - ydx)/(x^2+y^2)$, o "elemento de ângulo", definida no Capítulo IV. Sabemos que ω é uma forma fechada de classe C^∞, com $\int_{\partial M} \omega = \int_{S_1} \omega = 2\pi$, enquanto $\int_M d\omega = \int_M 0 = 0$. Portanto $\int_{\partial M} \omega \neq \int_M d\omega$.

9 Grau de uma aplicação

Sejam $f\colon M \to N$ um difeomorfsmo que preserva orientação e ω uma forma contínua de suporte compacto e grau máximo em N. Sabemos que a igualdade $\int_M f^*\omega = \int_N \omega$ é um enunciado global da fórmula de mudança de variáveis em integrais múltiplas. [Essencialmente, estamos usando f para introduzir novos parâmetros em N e $f^*\omega$ é a expressão de ω relativamente a tais parâmetros.] Sabemos também que $\int_M f^*\omega = -\int_N \omega$ se a mudança de variável é feita por um difeomorfismo f que inverta a orientação. Neste parágrafo, generalizaremos amplamente essas fórmulas de mudança de variável. Mostraremos que (pelo menos quando M é orientada, compacta, sem bordo e N é orientada, conexa, com dim. M = dim. $N = m$) dada *qualquer* aplicação $f\colon M \to N$ de classe C^1, existe um inteiro r, chamado o grau de f, tal que $\int_M f^*\omega = r \cdot \int_N \omega$, seja qual for a forma contínua de grau m, com suporte compacto, em N. Esse inteiro r, que depende apenas da classe de homotopia de f, ocupa posição de grande destaque em Topologia, em Análise e suas aplicações. (Veja [13], [15] e [16].)

Obteremos o grau examinando a imagem inversa de um valor regular de uma aplicação própria $f\colon M \to N$, de classe C^1, onde M e N são

superfícies orientadas de mesma dimensão.

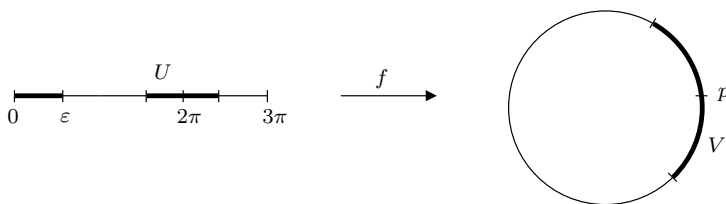
Dizer que $y \in N$ é valor regular de f significa, como sabemos, que para cada ponto $x \in M$ tal que $f(x) = y$, a derivada $f'(x)\colon T_xM \to T_yN$ é sobrejetiva. Como M e N têm a mesma dimensão, isto quer dizer que $f'(x)$ é um isomorfismo. Pelo Teorema da Aplicação Inversa, existem abertos $U \ni x$ em M e $V \ni y$ em N tais que f aplica U difeomorficamente sobre V. Em particular, x é o único ponto de U que é transformado por f em y, isto é, $f^{-1}(y) \cap U = \{x\}$, ou ainda, x é um ponto isolado de $f^{-1}(y)$. Assim, se M e N têm a mesma dimensão e $y \in N$ é um valor regular de uma aplicação $f\colon M \to N$, de classe C^1, então a imagem $f^{-1}(y)$ é um subconjunto (fechado e) discreto de M, isto é, todos os seus pontos são isolados.

Se, além disso, a aplicação $f\colon M \to N$ é própria (em particular, se M é compacta) então, como $\{y\}$ é compacto, a imagem inversa $f^{-1}(y)$ é compacta. Como nenhum dos seus pontos é ponto de acumulação, segue-se que $f^{-1}(y) = \{x_1, \ldots, x_k\}$ é um conjunto finito. Estas considerações são completadas pelo

Teorema 19. *Sejam M, N superfícies de mesma dimensão e $f\colon M \to N$ uma aplicação própria de classe C^1. Todo valor regular $y \in N$ possui uma vizinhança aberta V tal que $f^{-1}(V) = U_1 \cup \cdots \cup U_k$ é reunião finita de abertos U_i, dois a dois disjuntos, cada um dos quais é transformado por f difeomorficamente sobre V.*

Demonstração: Seja $f^{-1}(y) = \{x_1, \ldots, x_k\}$. Para cada $i = 1, \ldots, k$, existem $W_i \ni x_i$ aberto em M e $Z_i \ni y$ aberto em N tais que f transforma W_i difeomorficamente sobre Z_i. Seja c a menor distância entre dois pontos distintos quaisquer de $f^{-1}(y)$. Substituindo, se necessário, W_i por $W_i \cap B(x_i, c/2)$, podemos supor que os W_i são dois a dois disjuntos. O conjunto $F = M - \bigcup\limits_{i=1}^{k} W_i$ é fechado em M, logo sua imagem $f(F)$ é fechada em M (conforme o Lema 1 da seção 8). Como nenhum x_i pertence a F, temos $y \notin f(F)$. Assim, existe $V \ni y$, aberto em N, tal que $V \cap f(F) = \varnothing$. Substituindo, se necessário, V por $\left(\bigcap\limits_{i=1}^{k} Z_i\right) \cap V$, podemos ainda supor que $V \subset Z_1 \cap \cdots \cap Z_k$. A condição $V \cap f(F) = \varnothing$ significa que $f^{-1}(V) \subset W_1 \cup \cdots \cup W_k$. Ponhamos agora, para cada $i = 1, \ldots, k$, $U_i = f^{-1}(V) \cap W_i = (f|W_i)^{-1}(V)$. Como $f|W_i$ é um difeomorfismo de W_i sobre Z_i e $V \subset Z_i$, resulta que f transforma U_i difeomorficamente sobre V. Finalmente, $f^{-1}(V) = f^{-1}(V) \cap [W_1 \cup \cdots \cup W_k] =$

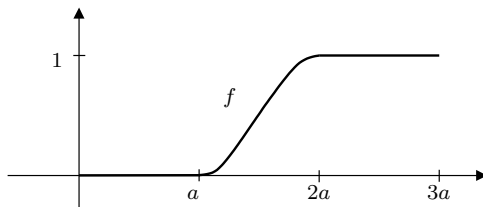
$[f^{-1}(V) \cap W_1] \cup \cdots \cup [f^{-1}(V) \cup W_k] = U_1 \cup \cdots \cup U_k\,.$

Exemplo 32. A hipótese de f ser própria não entra no Teorema 19 apenas para garantir que $f^{-1}(y)$ seja finito. O exemplo seguinte ilustra bem o que pode ocorrer sem essa hipótese. Seja $f\colon (0,3\pi) \to S^1$ definida por $f(t) = (\cos t, \operatorname{sen} t)$. Todo ponto $y \in S^1$ é valor regular de f. Em particular, tomemos $p = (1,0)$. Então $f^{-1}(p) = \{2\pi\}$. Entretanto, seja qual for o aberto V em S^1 contendo p, $f^{-1}(V)$ contém, além de uma vizinhança U do ponto 2π, também um intervalo do tipo $(0,\varepsilon)$, o qual não possui pontos de $f^{-1}(p)$ nem se aplica sobre V. [Como S^1 é compacto mas $(0,3\pi)$ não é, f não é própria.]

Por ser o caso mais simples, e mais útil, definiremos primeiro o grau de uma aplicação $f\colon M \to S^m$, onde M é uma superfície m-dimensional, compacta, orientada e sem bordo. Usaremos o

Lema. *Seja $p \in A \subset S^m$, onde A é aberto. Existem abertos W, V em S^m, com $p \in W \subset V \subset A$, e uma aplicação $\lambda\colon S^m \to S^m$, de classe C^∞, tal que*

1) *$\lambda(S^m - V) = \{q\}$, onde $q = -p$ é o antípoda de p;*
2) *$\lambda|V$ é um difeomorfismo de V sobre $S^m - \{q\}$.*
3) *$\lambda(x) = x$ para todo $x \in W$;*
4) *λ é C^∞-homotópica à aplicação identidade de S^m.*

Demonstração: Seja $\varphi\colon S^m - \{p\} \to \mathbb{R}^m$ a projeção estereográfica. Temos $\varphi(q) = 0$ e o compacto $\varphi(S^m - A)$ está contido numa bola $B_a = B(0;a)$. Pomos $V = S^m - \varphi^{-1}(\overline{B}_a)$ e $W = S^m - \varphi^{-1}(\overline{B}_{2a})$, onde $B_{2a} = B(0;2a)$. Como φ é um difeomorfismo, basta definir uma aplicação $\mu\colon \mathbb{R}^m \to \mathbb{R}^m$, de classe C^∞, tal que 1) $\mu(B_a) = \{0\}$; 2) μ aplica $\mathbb{R}^m - \overline{B}_a$ difeomorficamente sobre $\mathbb{R}^m - \{0\}$; 3) $\mu(x) = x$ para todo $x \in \mathbb{R}^m - \overline{B}_{2a}$; 4) Existe uma homotopia $H\colon \mu \simeq \mathrm{id}_{\mathbb{R}^m}$ tal que $H(t,x) = x$ para todo x com norma $\geq 2a$.

A definição de μ se faz com auxílio de uma função $f\colon \mathbb{R} \to \mathbb{R}$, de classe C^∞, tal que $f(t) = 0$ para $t \leq a$, $\quad f(t) = 1$ quando $t \geq 2a$ e

f, restrita ao intervalo $(a, 2a)$, é um difeomorfismo sobre $(0, 1)$. [Isto equivale a dizer que $a < t < 2a \Rightarrow f'(t) > 0$.] O gráfico de f é esboçado acima. Sua definição utiliza a função $\delta\colon \mathbb{R} \to \mathbb{R}$, introduzida no início do parágrafo 4: $f(t) = \delta(t/a - 3)$ se $t < 3a$ e $f(t) = 1$ se $t \geq 3a$. Então obtemos μ pondo $\mu(x) = f(|x|) \cdot x$ para todo $x \in \mathbb{R}^m$, onde $|x|$ é a norma euclidiana. Para nos convencermos de que μ é um difeomorfismo de $\mathbb{R}^m - \overline{B}_a$ sobre $\mathbb{R}^m - \{0\}$, observemos primeiro que $t \mapsto f(t) \cdot t$ é um difeomorfismo de $(a, +\infty)$ sobre $(0, +\infty)$; segundo, que $k(u, t) = (u, f(t)t)$ define um difeomorfismo de $S^{m-1} \times (a, +\infty)$ sobre $S^{m-1} \times (0, +\infty)$; terceiro, que as aplicações $h_0\colon S^{m-1} \times (0, +\infty) \to \mathbb{R}^m - \{0\}$ e $h_a\colon S^{m-1} \times (a, +\infty) \to \mathbb{R}^m - \overline{B}_a$, ambas definidas por $(u, t) \mapsto tu$, são difeomorfismos; quarto, que $\mu = h_0 \circ k \circ h_a^{-1}$. Finalmente, a homotopia $H\colon I \times \mathbb{R}^m \to \mathbb{R}^m$ é linear: $H(t, x) = (1 - t)\mu(x) + tx$.

Sejam M, N superfícies orientadas sem bordo, de mesma dimensão m, e $f\colon M \to N$ uma aplicação própria de classe C^1. Para todo valor regular $y \in N$, de f, sabemos que $f^{-1}(y) = \{x_1, \ldots, x_k\}$ é finito. Diremos que um ponto $x_i \in f^{-1}(y)$ é *positivo* ou *negativo*, conforme o isomorfismo $f'(x_i)\colon T_{x_i}M \to T_yN$ preserve ou inverta orientação.

O *grau de f no ponto y*, indicado por $gr_y f$, é, por definição, a diferença entre o número de pontos positivos e o número de pontos negativos em $f^{-1}(y)$. Segue-se imediatamente do Teorema 19 que existe uma vizinhança aberta V do valor regular y em N tal que todo $z \in V$ é valor regular de f e, além disso, $gr_z f = gr_y f$ para qualquer $z \in V$. No que segue, mostraremos que se N é conexa e f é de classe C^2 então $gr_y f$ tem o mesmo valor, seja qual for o valor regular $y \in N$.

Como dissemos acima, começaremos com o caso em que M é compacta e $N = S^m$. Temos então o

Teorema 20. *Seja M uma superfície compacta, m-dimensional, orientada e sem bordo. A toda aplicação $f\colon M \to S^m$, de classe C^2, corres-*

ponde um inteiro r tal que:

1) *Para todo valor regular y de f tem-se $gr_y f = r$;*

2) *Para toda m-forma contínua ω em S^m, tem-se $\int_M f^*\omega = r \cdot \int_{S^m} \omega$.*

Demonstração: Pelo Teorema de Sard, f possui um valor regular $p \in S^m$. Embora isto não seja estritamente necessário, consideraremos separadamente, e em primeiro lugar, o caso em que $f^{-1}(p) \neq \varnothing$. Pelo Teorema 19, existe uma vizinhança $V \ni p$, aberta em S^m, tal que $f^{-1}(V) = U_1 \cup \cdots \cup U_k$ é reunião de abertos dois a dois disjuntos, cada um dos quais se aplica difeomorficamente por f sobre V. Podemos supor que V é o aberto mencionado no lema acima. Seja $g = \lambda \circ f \colon M \to S^m$, onde $\lambda \colon S^m \to S^m$ é dada pelo referido lema. Como $\lambda \simeq \mathrm{id} \colon S^m \to S^m$, temos $g \simeq f$. Pelo Teorema 17, $\int_M f^*\omega = \int_M g^*\omega$ para toda m-forma ω, contínua em S^m. Ora, g é constante, igual a $q = -p$, no conjunto fechado $F = M - \bigcup_{i=1}^{k} U_i$, logo $g^*\omega = 0$ em todos os pontos de F. Pelo Teorema da Irrelevância dos Zeros, temos

$$\int_M g^*\omega = \int_{M-F} g^*\omega = \sum_{i=1}^{k} \int_{U_i} g^*\omega,$$

a última igualdade valendo porque $M - F$ é reunião disjunta dos abertos U_i. Restrito a cada U_i, g é um difeomorfismo sobre $S^m - \{q\}$. Portanto

$$\int_{U_i} g^*\omega = \pm \int_{S^m - \{q\}} \omega = \pm \int_{S^m} \omega,$$

a última igualdade valendo em virtude do Teorema da Irrelevância dos Conjuntos de Medida Nula. O sinal é $+$ ou $-$, conforme os pontos $x \in U_i$ sejam positivos ou negativos. Logo

$$\sum_{i=1}^{k} \int_{U_i} g^*\omega = gr_p f \cdot \int_{S^m} \omega, \text{ ou seja }, \int_M f^*\omega = gr_p f \cdot \int_{S^m} \omega.$$

Se $y \in S^m$ é qualquer outro valor regular de f, tomando uma m-forma contínua ω em S^m, com $\int_{S^m} \omega \neq 0$, o argumento acima nos dará $gr_y f = \int_M f^*\omega / \int_{S^m} \omega = gr_p f$, o que demonstra este caso. Se, porém, o valor regular $p \in S^m$ é tal que $f^{-1}(p) = \varnothing$ então $gr_p f = 0$. Este caso significa que $f(M) \subset S^m - \{p\}$. Como $S^m - \{p\}$ é difeomorfo a $\mathbb{R}^m$, que é contrátil, a aplicação f é, então homotópica a uma constante.

Pelo Teorema 17, obtemos então $\int_M f^*\omega = 0$, seja qual for a m-forma contínua ω em S^m. Pelo primeiro caso, teremos $gr_y f = \int_M f^*\omega / \int_{S^m} \omega = 0 = gr_p f$ para qualquer outro valor regular $y \in S^m$. Isto completa a demonstração.

O número $r = gr.f$, definido no Teorema 17, chama-se o *grau* da aplicação $f\colon M \to S^m$. Ele goza das seguintes propriedades:

1) *Se $f, g\colon M \to S^m$ são homotópicas em classe C^2 então $gr.f = gr.g$;*
2) *Se $f\colon M \to S^m$ não é sobrejetiva então $gr.f = 0$.*

A propriedade 1 resulta imediatamente da caracterização $gr.f = \int_M f^*\omega / \int_{S^m} \omega$, onde ω é uma (qualquer) m-forma contínua em S^m cuja integral não é zero, juntamente com o Teorema 17. A propriedade 2, está contida na demonstração acima.

Exemplo 33. Seja $M \subset \mathbb{R}^{m+1}$ uma hiperfície compacta orientada. Indiquemos com $K\colon M \to \mathbb{R}$ a função curvatura de Gauss-Kronecker de M (cfr. Exemplo 24) e com c_m o volume (área, se $m = 2$, comprimento, se $m = 1$) da esfera unitária S^m. O número $\int_M K$ chama-se a *curvatura integral* de M. Segue-se do que vimos no exemplo acima citado que $\int_M K = r \cdot c_m$, onde $r \in Z$ é o grau da aplicação normal de Gauss $\gamma\colon M \to S^m$. Quando m é par, demonstra-se que o inteiro r é um invariante topológico de M; mais precisamente, $2r$ é a "característica de Euler" da hiperfície M. (Veja [13], Capítulo IV.)

O grau tem sentido, mais geralmente, para qualquer aplicação *contínua* $f\colon M \to S^m$, onde M é uma superfície m-dimensional compacta, sem bordo. Com efeito, existe $g\colon M \to S^m$ de classe C^2, homotópica a f. Pomos, por definição, $gr.f = gr.g$. Não há ambigüidade porque se $h\colon M \to S^m$ é outra aplicação de classe C^2 homotópica a f então, por transitividade, h é homotópica a g, logo $h \simeq g$ em classe C^2, donde $gr.g = gr.h$. As propriedades 1. e 2. acima continuam válidas para o grau de uma aplicação contínua $f\colon M \to S^m$.

Exemplo 34. Segue-se do Teorema de Stokes que se $M = \partial N$ é o bordo de uma superfície $(m+1)$-dimensional orientada N e $f\colon M \to S^m$ é a restrição de uma aplicação, $F\colon N \to S^m$, de classe C^2, então $gr.f = 0$. Com efeito, tomemos em S^m uma m-forma contínua positiva ω. Temos $gr.f = \int_M f^*\omega / \int_{S^m} \omega$. Mas, como $d\omega = 0$, temos

$$\int_M f^*\omega = \int_{\partial N} F^*\omega = \int_N d(F^*\omega) = \int_N F^*(d\omega) = 0.$$

O mesmo resultado vale com "classe C^2"substituido por "contínua". Basta tomar uma aplicação $G\colon N \to S^m$, de classe C^2, homotópica a F e usar o mesmo argumento. Em particular, toda aplicação $f\colon S^m \to S^m$ que possui uma extensão contínua $F\colon B \to S^m$ (onde B é a bola fechada unitária de $\mathbb{R}^{m+1}$) tem grau zero. Este caso particular (que implica o Teorema 11) resulta também de f ser então homotópica a uma constante.

Por meio do teorema abaixo, estenderemos o conceito de grau de modo que ele tenha sentido para aplicações próprias entre duas superfícies orientadas de mesma dimensão.

Teorema 21. *Sejam M,N superfícies m-dimensionais orientadas, sem bordo. Se N é conexa, a cada aplicação própria $f\colon M \to N$, de classe C^2, corresponde um número inteiro r com as seguintes propriedades:*
1) *Para cda valor regular $y \in N$, de f, tem-se $gr_y f = r$;*
2) *Para toda m-forma contínua ω, integrável em N, tem-se $\int_M f^*\omega = r \cdot \int_N \omega$.*

A demonstração do Teorema 20 utilizará o "Lema de Isotopia", o qual será precedido de algumas definições.

Sejam $f, g\colon M \to N$ difeomorfismos. Uma *isotopia* entre f e g é uma homotopia $H\colon I \times M \to N$ entre f e g tal que, para todo $t \in I$, a aplicação $H_t\colon M \to N$, definida por $H_t(x) = H(t, x)$, é um difeomorfismo. Se f e g são difeomorfismos de M sobre si mesmo, diz-se que uma isotopia $H\colon f \simeq g$ tem *suporte compacto* quando existe um compacto $K \subset M$ tal que $H(x, t) = x$ para todo $x \in M - K$ e todo $t \in I$. Analogamente, diz-se que um difeomorfismo $f\colon M \to N$ tem suporte compacto quando existe um compacto $K \subset M$ tal que $f(x) = x$ para todo $x \in M - K$.

Lema de Isotopia. *Seja M uma superfície conexa de classe C^k. Dados dois pontos $a, b \in M$, existe um difeomorfismo $f\colon M \to N$, tal que $f(a) = b$, isotópico à identidade por uma isotopia de classe C^k com suporte compacto.*

Demonstração: O Lema será provado em três etapas: a, b e c.

a) A superfície M é o espaço euclidiano $\mathbb{R}^m$. Então tomaremos a função auxiliar $\xi\colon \mathbb{R}^m \to \mathbb{R}$, de classe C^∞, definida no início do parágrafo 4, uma constante $c > 0$ tal que $c \cdot |b-a| \cdot \sup\limits_{y \in \mathbb{R}^m} |\xi'(y)| < 1$ e poremos $f(x) = x + \xi(c(x-a)) \cdot (b-a)$. Vemos que f é a soma da aplicação identidade com uma contração. Pelo Teorema da Perturbação da Identidade, concluimos que $f\colon \mathbb{R}^m \to \mathbb{R}^m$ é um difeomorfismo de classe C^∞. Como $\xi(c(x-a)) =$

0 se $c|x-a| > 2$, temos $f(x) = x$ fora da bola de centro a e raio $2/c$. Portanto f tem suporte compacto. Para todo $t \in [0,1]$, a aplicação $f_t \colon \mathbb{R}^m \to \mathbb{R}^m$, definida por $f_t(x) = x + t \cdot \xi(c(x-a)) \cdot (b-a)$ é uma perturbação da identidade, donde é um difeomorfismo de $\mathbb{R}^m$. Pondo $H(t,x) = f_t(x)$, obtemos uma isotopia de suporte compacto e classe C^∞ entre f e a aplicação identidade de $\mathbb{R}^m$.

b) Em seguida, mostraremos que o lema vale localmente em M, isto é, todo ponto $p \in M$ possui uma vizinhança aberta V tal que, dados arbitrariamente $a, b \in V$, existe um difeomorfismo $f \colon M \to M$, de classe C^k, com $f(a) = b$, isotópico à aplicação identidade de M por uma isotopia de classe C^k e suporte compacto. Para isto basta tomar uma parametrização $\varphi \colon \mathbb{R}^m \to V \subset M$, de classe C^k, com $\varphi(0) = p$. Dados $a = \varphi(a_0)$ e $b = \varphi(b_0)$ em V, seja $f_0 \colon \mathbb{R}^m \to \mathbb{R}^m$ o difeomorfismo C^∞-isotópico à identidade, com $f_0(a_0) = b_0$, obtido como se viu acima. Ponhamos $f = \varphi \circ f_0 \circ \varphi^{-1} \colon M \to M$. As verificações são imediatas.

c) Finalmente, dado $a \in M$, seja X o conjunto dos pontos $b \in M$ tais que existe um difeomorfismo $f \colon M \to M$, nas condições impostas pelo enunciado do Lema. O conjunto X não é vazio, pois evidentemente $a \in X$. Afirmamos que X é aberto em M. Com efeito, sejam $b \in X$ e $f \colon M \to M$ o difeomorfismo com $f(a) = b$, nas condições do enunciado do Lema. Pela etapa anterior, existe uma vizinhança $V \ni b$, aberta em M tal que, para cada $c \in V$ pode-se obter um difeomorfismo $g \colon M \to M$, de classe C^k, compactamente isotópico à identidade, com $g(b) = c$. Então $h = g \circ f \colon M \to M$ é um difeomorfismo de classe C^k, também compactamente isotópico à identidade, com $h(a) = c$. Portanto $c \in X$, ou seja, $V \subset X$. De modo análogo se mostra que $M - X$ é aberto. Pela conexidade de M, concluimos que $X = M$, o que prova o Lema.

Demonstração do Teorema 21: Mostremos primeiramente que $gr_y f$ não depende do valor regular $y \in N$. Com efeito, se $V \ni y$ é uma vizinhança (conexa) em N tal que $f^{-1}(V) = \bigcup_{i=1}^{k} U_i$ é reunião disjunta de abertos $U_i \subset M$, cada um dos quais se aplica por f difeomorficamente sobre V (Teorema 19) e ω é uma m-forma contínua com $\operatorname{supp.}\omega \subset V$ e $\int_N \omega \neq 0$ então, para cada $i = 1, \ldots, k$, $\int_{U_i} f^*\omega = \pm \int_V \omega$ (sinal $+$ ou $-$ conforme $f|U_i$ preserve ou inverta orientação). Como $\operatorname{supp.}(f^*\omega) \subset$

$\bigcup_{i=1}^{k} U_i$, temos

$$\int_M f^*\omega = \sum_{i=1}^{k} \int_{U_i} f^*\omega = \sum_{i=1}^{k} \pm \int_V \omega = (gr_y f) \cdot \int_N \omega.$$

Seja agora $z \in N$ outro qualquer valor regular de f. Pelo Lema de Isotopia, existe um difeomorfismo $h\colon N \to N$, compactamente isotópico à identidade em classe C^2, tal que $h(z) = y$. Seja $W = h^{-1}(V)$. Restringindo V, se for necessário, podemos supor também que $f^{-1}(W)$ é reunião disjunta de abertos que se aplicam por f difeomorficamente sobre W. Então, como a forma $h^*\omega$ tem suporte contido em W e $\int_N h^*\omega = \int_N \omega \neq 0$, o raciocínio acima feito para ω e y nos permite escrever

$$gr_z f = \int_M f^*(h^*\omega) \Big/ \int_N h^*\omega = \int_M (h \circ f)^*\omega \Big/ \int_N h^*\omega = \\ = \int_M f^*\omega \Big/ \int_N \omega = gr_y f.$$

Nas igualdades acima, fizemos uso do fato de que, sendo h homotópico à identidade (e conseqüentemente $h \circ f \simeq f$), temos $\int_M (h \circ f)^*\omega = \int_M f^*\omega$ e $\int_N h^*\omega = \int_N \omega$.

Seja então $r = gr_y f$, onde $y \in N$ é qualquer valor regular de f. Provaremos agora que $\int_M f^*\omega = r \cdot \int_N \omega$ para todo m-forma contínua ω, integrável em N. Dividiremos a demonstração em três casos.

Caso 1. A aplicação própria $f\colon M \to N$ é um difeomorfismo local que preserva orientação. Seja uma forma contínua integrável $\omega \geq 0$ em N. Como todo ponto $y \in N$ é um valor regular, podemos tomar uma cobertura aberta localmente finita $N = \cup V_i$ tal que cada V_i tem fecho compacto em N e $f^{-1}(V_i)$ é uma reunião finita de r abertos, dois a dois disjuntos, cada um dos quais se aplica por f difeomorficamente sobre V_i. Escrevamos $\omega = \Sigma\, \omega_i$, onde cada ω_i é ≥ 0, contínua e $\operatorname{supp.} \omega_i \subset V_i$. Então $M = \cup f^{-1}(V_i)$ é uma cobertura localmente finita. [Mostra-se facilmente que se $f\colon X \to Y$ é contínua e $\mathcal{C}$ é uma família localmente finita em Y então $f^{-1}(\mathcal{C})$ é localmente finita em X.] Para cada i, $f^*\omega_i$ é uma forma ≥ 0, com $\operatorname{supp.} f^*\omega_i \subset f^{-1}(V_i)$, $f^{-1}(\overline{V}_i)$ compacto e, como já vimos acima, $\int_M f^*\omega_i = r \cdot \int_N \omega_i$. Temos

$$f^*\omega = \sum_i f^*\omega_i \text{ e } \int_M f^*\omega = \sum_i \int_M f^*\omega_i = r \cdot \sum_i \int_N \omega_i = r \cdot \int_N \omega.$$

[Em particular, ω integrável $\Rightarrow f^*\omega$ integrável.] Se a forma integrável ω não é ≥ 0, podemos escrevê-la como $\omega = \alpha - \beta$, com α, β integráveis e ≥ 0. Temos ainda $\int_M f^*\omega = \int_M f^*\alpha - \int_M f^*\beta = r \cdot \int_N \alpha - r \cdot \int_N \beta = r \cdot \int_N \omega$. Evidentemente, o mesmo argumento se aplica quando f inverte orientação.

Caso 2. A aplicação própria $f\colon M \to N$ é um difeomorfismo local, que não supomos preservar nem inverter orientação. [Somente quando M é desconexa este caso é mais geral do que o anterior.] Então $M = A \cup B$, onde A é o conjunto dos pontos $x \in M$, tais que $f'(x)$ preserva orientação e $B = M - A$. Ambos, A e B, são abertos e as restrições $f|A\colon A \to N$, $f|B\colon B \to N$ são próprias. [Por exemplo, se $x_k \to \infty$ em A então $x_k \to \infty$ em M, pois B é aberto em M, logo $f(x_k) \to \infty$ em N devido a $f\colon M \to N$ ser própria.] Temos $r = r_A + r_B$, onde $r_A = gr_y(f|A)$ e $r_B = gr_y(f|B)$ para todo $y \in N$. Pelo Caso 1, vem

$$\int_M f^*\omega = \int_A f^*\omega + \int_B f^*\omega = \int_A (f|A)^*\omega + \int_B (f|B)^*\omega =$$
$$= r_A \cdot \int_N \omega + r_B \cdot \int_N \omega = r \cdot \int_N \omega.$$

Observação: Nos casos acima, não é estritamente necessário que N seja conexa. Esta hipótese entra apenas para garantir que $gr_y f = r$ (constante) para todo $y \in N$. Se soubermos disto por outras razões teremos $\int_M f^*\omega = r \cdot \int_N \omega$ sejam quais forem a forma contínua integrável ω em N e o difeomorfismo local $f\colon M \to N$.

Caso 3. Em que $f\colon M \to N$ é qualquer aplicação própria de classe C^2. Seja S o subconjunto fechado de M formado pelos pontos x nos quais a derivada $f'(x)$ não é um isomorfismo. Em todos os pontos $x \in S$ temos $(f^*\omega)(x) = 0$. Como f é própria, $f(S)$ é um subconjunto fechado de N o qual, pelo Teorema de Sard, tem medida nula. Como os zeros do integrando e os conjuntos de medida nula são irrelevantes na integração temos

$$\int_M f^*\omega = \int_{M-S} f^*\omega \quad \text{e} \quad \int_N \omega = \int_{N-f(S)} \omega.$$

Observe que $f^{-1}(f(S)) = S \cup T$, onde T é o conjunto dos pontos $x \in M$ tais que $f(x) \in f(S)$ e $f'(x)$ é um isomorfismo, ou seja $T = f^{-1}(f(S)) \cap (M - S)$. Assim, T é um subconjunto fechado da superfície $M - S$. Todo ponto $x \in T$ possui uma vizinhança U em M que é aplicada difeomorficamente por f sobre uma vizinhança V do ponto $y = f(x)$ em

N. Evidentemente $f(U \cap T) \subset V \cap f(S)$. Como $V \cap f(S)$ tem medida nula, o mesmo ocorre com $U \cap T$. Como $x \in T$ é qualquer, segue-se que T é um conjunto de medida nula em $M - S$. Portanto

$$\int_M f^*\omega = \int_{M-S} f^*\omega = \int_{(M-S)-T} f^*\omega = \int_{M-(S\cup T)} f^*\omega.$$

Ora, a aplicação $g\colon M-(S\cup T) \to N-f(S)$, obtida por restrição de f, é um homeomorfismo local, com $gr_y g = gr_y f = r$ para todo $y \in N-f(S)$. Resulta do caso anterior que, para toda m-forma contínua integrável ω sobre N, vale

$$\int_M f^*\omega = \int_{M-(S\cup T)} f^*\omega = \int_{M-(S\cup T)} g^*\omega = r \cdot \int_{N-f(S)} \omega = r \cdot \int_N \omega,$$

o que encerra a demonstração.

O grau de uma aplicação própria $f\colon M \to N$, de classe C^2, entre superfícies orientadas de mesma dimensão, (N conexa) goza das seguintes propriedades:

1) *Se $f\colon M \to M$ é a aplicação identidade, então $gr.f = 1$;*

2) *Se $f, g\colon M \to N$ são propriamente homotópicas em classe C^2 então $gr.f = gr.g$.*

Com efeito, seja ω uma m-forma contínua em N, com suporte compacto, tal que $\int_N \omega \neq 0$. Pelo Teorema 17a, temos

$$gr.f = \int_M f^*\omega \Big/ \int_N \omega = \int_M g^*\omega \Big/ \int_N \omega = gr.g;$$

3) *Se $gr.f \neq 0$ então f é sobrejetiva.*

Com efeito, se existir $y \in N - f(M)$ então, por definição, y será valor regular de f. Como $f^{-1}(y) = \varnothing$, teremos $gr_y f = 0$, logo $gr\, f = 0$.

4) *Se $f\colon M \to N$ e $g\colon N \to P$ são próprias, de classe C^2 M, N, P são orientadas, de mesma dimensão e N, P são conexas, então $gr(g \circ f) = (gr.g) \cdot (gr.f)$.*

Quando $gr.g \neq 0$, a propriedade acima pode ser demonstrada assim: tomemos uma m-forma contínua ω em P, com $\int_P \omega \neq 0$. Então $\int_N g^*\omega = (gr.g) \cdot \int_P \omega \neq 0$, logo

$$gr(g \circ f) = \frac{\int_M (g\circ f)^*\omega}{\int_P \omega} = \frac{\int_M f^*(g^*\omega)}{\int_N g^*\omega} \cdot \frac{\int_N g^*\omega}{\int_P \omega} = (gr.g) \cdot (gr.f).$$

No caso geral, existe, pelo Teorema de Sard, um ponto $z \in P$ que é valor regular de g e de $g \circ f$. Então $g^{-1}(z) = \{y_1, \ldots, y_k\}$, onde cada y_i é valor regular de f. Seja $f^{-1}(y_i) = \{x_{i1}, \ldots, x_{is_i}\}$, donde $(g \circ f)^{-1}(z) = \{x_{ij}; i = 1, \ldots, k,\ j = 1, \ldots, s_i\}$. Ponhamos $\varepsilon_i = 1$ se o ponto y_i é positivo e $\varepsilon_i = -1$ no caso contrário. Diremos que ε_i é o *sinal* do ponto y_i em relação à aplicação g. Analogamente, sejam η_{ij} o sinal do ponto x_{ij} em relação a f e ρ_{ij} o sinal de x_{ij} em relação a $g \circ f$. Temos $\rho_{ij} = \eta_{ij} \cdot \varepsilon_i$, $\Sigma\, \varepsilon_i = gr_y g$, $\sum\limits_{i,j} \rho_{ij} = gr_z(g \circ f)$ e $\sum\limits_j \eta_{ij} = gr_y f = gr.f$ (para todo i). Logo

$$gr(g \circ f) = gr_z(g \circ f) = \sum_{i,j} \rho_{ij} = \sum_i \Big(\sum_j \eta_{ij}\Big)\varepsilon_i = \sum_i (gr.f)\varepsilon_i =$$
$$= (gr.f) \cdot \sum_i \varepsilon_i = (gr.f) \cdot (gr.g).$$

A extensão da noção de grau para aplicações próprias entre superfícies não-compactas permite que a utilizemos no caso em que tais superfícies se reduzem a abertos do espaço euclidiano, como no exemplo seguinte.

Exemplo 35. Sejam $U, V \subset \mathbb{R}^m$ abertos (V conexo) e $f \colon U \to V$ uma aplicação própria de classe C^2. Suponhamos que o determinante jacobiano de f não mude de sinal em U (embora possa anular-se em alguns pontos). Se esse jacobiano é identicamente nulo então $f(U)$ tem medida nula, pelo Teorema de Sard. No caso contrário, existe um ponto $x \in U$ onde a derivada $f'(x) \colon \mathbb{R}^m \to \mathbb{R}^m$ é um isomorfismo, logo há um aberto $W \ni x$ que é aplicado por f difeomorficamente sobre um aberto $Z \ni f(x)$. Pelo Teorema de Sard, algum ponto $z \in Z$ é valor regular de f. Como $Z = f(W)$, a imagem inversa $f^{-1}(z)$ não é vazia. Todos os pontos de $f^{-1}(z)$ têm, por hipótese, o mesmo sinal. Logo $gr_z f \neq 0$, ou seja, $gr.f \neq 0$. Em particular, f aplica U sobre V. O mesmo argumento mostra que se M e N são superfícies orientadas de mesma dimensão, N conexa, e $f \colon M \to N$ é uma aplicação própria de classe C^2 cujo determinante jacobiano não muda de sinal em M então ou $f(M)$ tem medida nula em N (no caso em que o jacobiano é identicamente nulo) ou f é sobrejetiva. Em particular, concluimos que se a aplicação própria $f \colon M \to N$ é injetiva então ela também é sobrejetiva (dentro das demais hipóteses, inclusive N conexa). Com efeito, a injetividade de f impede que seu jacobiano mude de sinal: se $x_1, x_2 \in M$ fossem

tais que $f'(x_1) > 0$ e $f'(x_2) < 0$ então, pondo $y_1 = f(x_1)$ e $y_2 = f(x_2)$ teríamos (pela injetividade de f) $f^{-1}(y_1) = \{x_1\}$ e $f^{-1}(y_2) = \{x_2\}$ logo $gr_{y_1} f = 1$ e $gr_{y_2} f = -1$, uma contradição.

Tendo em vista principalmente as aplicações, é importante observar que, se M e N são superfícies orientadas de mesma dimensão, pode-se definir o grau de qualquer aplicação própria $f\colon M \to N$, mesmo que se tenha apenas $f \in C^0$. Isto é feito com base nos dois lemas seguintes:

1º) Se M e N são superfícies de classe C^k, toda aplicação própria $f\colon M \to N$, de classe C^0, é propriamente homotópica a uma aplicação $g\colon M \to N$, de classe C^k. Mais precisamente, dada N, existe uma função contínua positiva $\varepsilon\colon N \to \mathbb{R}$ tal que se $g\colon M \to N$ é tal que $|f(x)-g(x)| < \varepsilon(x)$ para todo $x \in M$ então g é propriamente homotópica a f.

2º) Se existe uma homotopia própria de classe C^0 entre duas aplicações $f, g\colon M \to N$ de classe C^k, existe também uma homotopia própria de classe C^k entre f e g.

No caso em que $N = S^m$, estes lemas decorrem da observação que fizemos no parágrafo 8, logo em seguida à demonstração do Teorema 17. Não provaremos o caso geral, para não nos estendermos em questões de Topologia Diferencial. (Veja [14].)

Admitidos os lemas, definiremos o grau da aplicação própria $f\colon M \to N$ pondo $gr(f) = gr(g)$, onde $g\colon M \to N$ é qualquer aplicação de classe C^2, propriamente homotópica a f. Se $h\colon M \to N$ é outra aplicação de classe C^2 nas mesmas condições, então existe uma homotopia própria $H\colon g \simeq h$, de classe C^2, logo $gr(g) = gr(h)$. Portanto $gr(f)$ está bem definido.

Usando novamente os lemas acima, prova-se que se $f, g\colon M \to N$ são propriamente homotópicas (em classe C^0) então $gr(f) = gr(g)$. Além disso, se $gr(f) \neq 0$ então a aplicação própria $f\colon M \to N$ (de classe C^0) é sobrejetiva. [Vejamos a demonstração deste último fato. Suponhamos, por absurdo, que exista $y \in N - f(M)$. Como f é própria, $f(M)$ é um subconjunto fechado de N, logo existe $\delta > 0$ tal que $|y - f(x)| \geq \delta$ para todo $x \in M$. Seja $\varepsilon\colon N \to \mathbb{R}$ a função contínua positiva mencionada no primeiro lema acima. Substituindo, se necessário, $\varepsilon(x)$ por $\min\{\varepsilon(x), \delta\}$, podemos supor que $\varepsilon(x) \leq \delta$ para todo $x \in M$. Portanto, se $g\colon M \to N$ é uma aplicação de classe C^2 tal que $|g(x) - f(x)| < \varepsilon(x)$ pra todo $x \in M$, teremos g propriamente homotópica a f, logo $gr(f) = gr(g)$. Teremos também $y \notin g(M)$, logo $gr(g) = 0$, e portanto $gr(f) = 0$.]

10 A integral de Kronecker

A seguir estenderemos, para um espaço euclidiano de dimensão qualquer, a fórmula devida a Kronecker e provada na seção 10 do Capítulo IV, que exprime o número "algébrico" de zeros de uma aplicação $F\colon D \to \mathbb{R}^2$, de classe C^2 num disco fechado do plano. Mesmo em dimensão 2, o resultado que obteremos agora será mais geral, pois se aplicará quando D for qualquer domínio compacto com fronteira regular, de classe C^2, no plano.

Seja, portanto, $D \subset \mathbb{R}^{m+1}$ uma superfície de dimensão $m+1$ e classe C^2, compacta, com bordo ("domínio compacto com fronteira regular" de classe C^2 em $\mathbb{R}^{m+1}$). O bordo ∂D é uma hiperfície orientável compacta em $\mathbb{R}^{m+1}$. Seja ainda $F\colon D \to \mathbb{R}^{m+1}$ uma aplicação de classe C^2 tal que $F(\partial D) \subset \mathbb{R}^{m+1} - \{0\}$. Indiquemos com $f_1, \ldots, f_{m+1}\colon D \to \mathbb{R}$ as funções-coordenada de F.

Um ponto $x \in D$ chama-se um *zero* de F quando $F(x) = 0$. Se o determinante jacobiano de F no ponto x é $\neq 0$, x chama-se um zero simples. Quando det. $Jf(x) > 0$ o zero simples x diz-se positivo e se det. $Jf(x) < 0$, x chama-se um zero negativo. Todo zero simples x é isolado, isto é, $F(y) \neq 0$ para todo y suficientemente próximo de x. Como D é compacto, se todos os zeros de F em D são simples, o conjunto desses zeros é finito.

Como no Capítulo IV, §10, nossa intenção é estimar o número de zeros de F em D. Para isso, introduziremos a *integral de Kronecker*

$$n(F; D) = \frac{1}{c_m} \int_{\partial D} \sum_{i=1}^{m+1} (-1)^{i+1} \frac{f_i \cdot df_1 \wedge \cdots \wedge \widehat{df_i} \wedge \cdots \wedge df_{m+1}}{|F|^{m+1}}$$

onde $c_m =$ volume da esfera unitária S^m e $|F(x)| = \left[\Sigma\, f_i(x)^2\right]^{1/2}$. Como $F(x) \neq 0$ para todo $x \in \partial D$, o integrando acima é uma forma de classe C^1 em ∂D. Estamos considerando em D a orientação de $\mathbb{R}^{m+1}$ e em ∂D a orientação induzida.

A integral de Kronecker goza das seguintes propriedades:

K1) *$n(F; D)$ é um número inteiro.*

Com efeito, sejam ω a forma elemento de volume de S^m, $\rho\colon \mathbb{R}^{m+1} - \{0\} \to S^m$ a projeção radial e $f = F|\partial D$. Então (vide Exemplo 17, no §2):

$$n(F; D) = \frac{1}{c_m} \int_{\partial D} f^* \rho^* \omega = \int_{\partial D} (\rho \circ f)^* \omega \Big/ \int_{S^m} \omega = gr.(\rho \circ f).$$

K2) *Se* $H\colon I\times\partial D\to\mathbb{R}^{m+1}$ *é uma homotopia de classe* C^2 *entre* $F|\partial D$ *e* $G|\partial D$*, tal que* $H(t,x)\neq 0$ *para todo* $t\in I$ *e todo* $x\in\partial D$ *então* $n(F;D)=n(G;D)$.

Com efeito, pondo $f=F|\partial D$ e $G=G|\partial D$ vemos que $L\colon I\times\partial D\to S^m$, definida por $L(t,x)=\rho(H(t,x))$, é uma homotopia de classe C^2 entre $\rho\circ f$ e $\rho\circ g$, logo $n(F;D)=gr(\rho\circ f)=gr(\rho\circ g)=n(G;D)$.

Com corolário de K2) temos o

Teorema de Rouché. *Sejam* $F,G\colon D\to\mathbb{R}^{m+1}$ *de classe* C^2*, com* $F(x)\neq 0$ *e* $|G(x)|<|F(x)|$ *para todo* $x\in\partial D$*. Então* $n(F+G;D)=n(F;D)$.

Com efeito, nestas condições a homotopia linear $H\colon I\times\partial D\to\mathbb{R}^{m+1}$, $H(t,x)=F(x)+t\cdot G(x)$, entre F e $F+G$, é tal que $H(t,x)\neq 0$ para todo $t\in I$ e todo $x\in\partial D$.

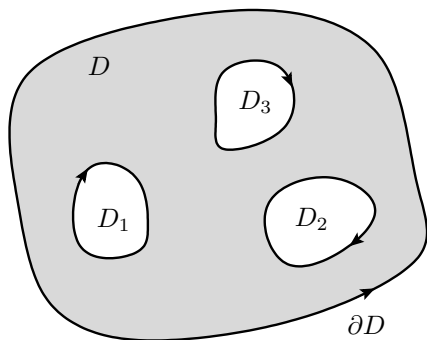
K3) *Se* $f(x)\neq 0$ *para todo* $x\in D$ *então* $n(F;D)=0$.

Com efeito, neste caso $\rho\circ F\colon D\to S^m$ tem sentido e é uma extensão de $\rho\circ f\colon\partial D\to S^m$. Segue-se que $n(F;D)=gr(\rho\circ f)=0$, conforme o Exemplo 34.

K4) *Sejam* $D_1,\dots,D_k\subset\operatorname{int.}D$ *domínios compactos com fronteiras regulares de classe* C^2*, dois a dois disjuntos. Se*

$$F\left(D-\bigcup_{i=1}^{k}\operatorname{int.}D_i\right)\subset\mathbb{R}^{m+1}-\{0\}\ \textit{então}\ n(F;D)=\sum_{i=1}^{k}n(F;D_i).$$

Observe que K3) é um caso particular de K4), com $k=1$ e $D_1=\varnothing$.

Para provar K4), observemos que $M=D-\bigcup\limits_{i=1}^{k}\operatorname{int.}D_i$ é um domínio compacto com fronteira regular. Considerando em M, D, $D_1,\dots,D_i$ a orientação de $\mathbb{R}^{m+1}$ e nos seus bordos as orientações induzidas, temos $\partial M=\partial D\cup(-\partial D_1)\cup\cdots\cup(-\partial D_k)$.

Indicaremos ainda com F todas as restrições desta aplicação. Como $\rho \circ F$ tem sentido em M, segue-se do Exemplo 34 que

$$0 = \int_{\partial M} (\rho \circ F)^* \omega = \int_{\partial D} (\rho \circ F)^* \omega - \sum_{i=1}^{k} \int_{\partial D_i} (\rho \circ F)^* \omega =$$

$$= n(F; D) - \sum_{i=1}^{k} n(F; D_i).$$

K5) *Se os zeros de F em D são todos simples então $n(F; D)$ é a diferença entre o número de zeros positivos e o número de zeros negativos.*

Sejam $x_1, \dots, x_k$ os zeros de F em D e $B_1 \ni x_1, \dots, B_k \ni x_k$ bolas fechadas duas a duas disjuntas, com centros nesses pontos. Se tomarmos os raios dessas bolas suficientemente pequenos, teremos, para cada $i = 1, \dots, k$, $F(x) = F'(x_i) \cdot (x - x_i) + r_i(x)$, com $|r_i(x)| < |F'(x_i) \cdot (x - x_i)|$ quando $x \in B_i$. Por K4), temos $n(F; D) = \sum\limits_{i=1}^{k} n(F; B_i)$ e, pelo Teorema de Rouché, $n(F; B_i) = n(T_i; B_i)$, onde $T_i(x) = F'(x_i) \cdot (x - x_i)$. Resta portanto mostrar que $n(T_i; B_i) = \pm 1$ conforme o sinal do determinante jacobiano det. $F'(x_i)$. Se este determinante for positivo, (vide Apêndice a este parágrafo) existe um caminho $A \colon I \to \mathcal{L}(\mathbb{R}^{m+1})$ tal que cada $A(t)$, $t \in I$, é uma transformação linear invertível, com $A(0) = F'(x_i)$ e $A(1) =$ identidade. Então $H \colon I \times B_i \to \mathbb{R}^{m+1}$, definida por $H(t, x) = A(t)(x - x_i)$, é uma homotopia entre T_i e a aplicação $J_i(x) = x - x_i$, onde $H(t, x) \neq 0$ para todo $x \neq x_i$. Logo $n(T_i; B_i) = n(J_i; B_i)$. Ora, evidentemente, $n(J_i; B_i) = 1$. Analogamente, se det. $F'(x_i) < 0$ obteremos uma homotopia entre T_i e a aplicação $L_i(x) = \bar{x} - x_i$,

onde, para $x = (x_1, \ldots, x_{n+1})$, pomos $\bar{x} = (-x_1, x_2, \ldots, x_{n+1})$. Como $n(L_i; B_i) = -1$, concluimos que $n(T_i; x_i) = \pm 1$ conforme o sinal de det. $F'(x_i)$, e isto completa a demonstração de K5).

A proposição K5) explica a notação: $n(F; D)$ é o "número algébrico" de zeros, pelo menos no caso em que esses zeros são todos simples. Observemos que esta última condição significa que $0 \in \mathbb{R}^{m+1}$ é um valor regular da aplicação F. Assim, K5) pode também ser expressa pela igualdade $n(F; D) = gr_0 F$. Nesta igualdade, o primeiro membro é um número determinado pelo comportamento de F no bordo ∂D, logo é um invariante global e m-dimensional. Enquanto isso, o segundo membro é um invariante local e $(m+1)$-dimensional de F, definido pela atuação de F na vizinhança dos seus zeros. O segundo membro teria sentido, mesmo que ∂D não fosse uma superfície, isto é, mesmo que o domínio compacto D não tivesse fronteira regular.

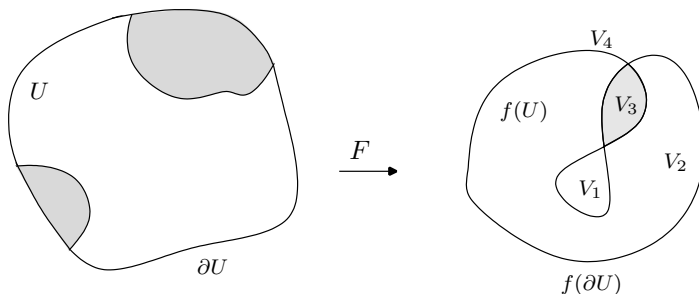
A última observação acima motiva a extensão do índice de Kronecker $n(F; D)$ para casos em que D é o fecho de qualquer aberto limitado do espaço euclidiano, originando uma outra maneira de apresentar a "teoria do grau", iniciada por Leray e Schauder e preferida dos especialistas em Análise Funcional Não-Linear. (Vide [19] e [15].) Esboçaremos a seguir essa abordagem do grau e mostraremos como reduzí-la ao que já sabemos.

Consideremos um aberto limitado $U \subset \mathbb{R}^m$ e uma aplicação contínua $F\colon \overline{U} \to \mathbb{R}$, cuja restrição $f = F|U$ é de classe C^2. Embora a fronteira de U em $\mathbb{R}^m$ não seja necessariamente uma superfície, vamos indicá-la com ∂U. Observemos que $\overline{U}$ e ∂U são compactos.

Para todo $y \in \mathbb{R}^m - F(\partial U)$ que seja valor regular de $f = F|U$, definiremos o *grau de F no ponto y*, que indicaremos com a notação $gr(F, U, y)$, do seguinte modo.

A aplicação $f\colon U - F^{-1}(F(\partial U)) \to \mathbb{R}^m - F(\partial U)$, de classe C^2, obtida por restrição de F, é própria, como resulta imediatamente da compacidade de $\overline{U}$. Analogamente, se V é qualquer componente conexa de $\mathbb{R}^m - F(\partial U)$, a restrição de f a $f^{-1}(V)$ é também uma aplicação própria $f_V\colon f^{-1}(V) \to V$. Sendo V conexa, podemos falar no grau de f_V. Então, para todo $y \in \mathbb{R}^m - F(\partial U)$, valor regular de $f = F|U$, poremos, por definição

$$gr(F, U, y) = gr_y\, f_V\,,$$

onde V é a componente conexa de $\mathbb{R}^m - F(\partial U)$ que contém o ponto y.

Com esta definição, é imediato que $gr(F,U,y)$ se mantém constante quando y varia numa componente conexa de $\mathbb{R}^m - F(\partial U)$. [Na figura acima, $\mathbb{R}^2 - F(\partial U)$ tem quatro componentes conexas, V_1, V_2, V_3 e V_4, sendo $gr(F,U,y)$ igual a zero se $y \in V_1$ ou $y \in V_4$, igual a 1 se $y \in V_2$ e igual a 2 se $y \in V_3$.]

Valem as seguintes propriedades:

1. *Se* $gr(F,U,y) \neq 0$ *então* $y \in F(U)$ *ou, mais precisamente,* $y \in \text{int.}\, f(U)$*;*

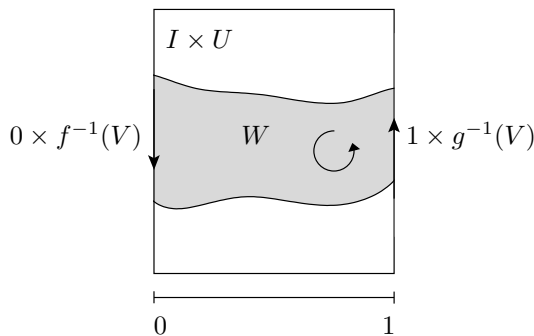
2. *Se existe uma homotopia* $H\colon I \times \overline{U} \to \mathbb{R}^m$ *de classe* C^2 *entre* F *e* G*, tal que* $y \notin H(I \times \partial U)$*, então* $gr(F,U,y) = gr(G,U,y)$*;*

3. *Se* $F = G$ *em* ∂U *então* $gr(F,U,y) = gr(G,U,y)$.

Em 2. e 3. acima, subentende-se "para todo $y \in \mathbb{R}^m - [F(\partial U) \cup G(\partial U)]$ que seja valor regular de f e g". Basta provar 2., pois 1. é evidente e 3. segue-se de 2. mediante a homotopia $H\colon I \times \overline{U} \to \mathbb{R}^m$, definida por $H(t,x) = (1-t)F(x) + t \cdot G(x)$.

Para a demonstração de 2., tomamos V aberto e conexo, tal que $y \in V \subset \mathbb{R}^m - H(I \times \partial U)$ e uma m-forma contínua ω, como suporte compacto, contido em V tal que $\int_V \omega \neq 0$. Então

$$gr(F,U,y) = \int_{f^{-1}(V)} f^*\omega \Big/ \int_V \omega \text{ e } \int_{G^{-1}(V)} g^*\omega \Big/ \int_V \omega.$$

Ora, o aberto $W = H^{-1}(V) \subset I \times U$ é uma superfície com bordo, sendo $\partial W = [1 \times g^{-1}(V)] \cup [-(0 \times f^{-1}(V))]$.

Portanto

$$0 = \int_W H^*(d\omega) = \int_W dH^*\omega = \int_{\partial W} H^*\omega =$$
$$= \int_{g^{-1}(V)} g^*\omega - \int_{f^{-1}(V)} f^*\omega.$$

Segue-se que $gr(G, U, y) = gr(F, U, y)$.

É importante mostrar que o grau de Leray-Schauder $gr(F, U, y)$ tem sentido para qualquer aplicação contínua $F\colon \overline{U} \to \mathbb{R}^m$, definida no fecho de um aberto limitado $U \subset \mathbb{R}^m$, e para qualquer ponto $y \in \mathbb{R}^m - F(\partial U)$. [Em particular, se $F|U \in C^2$, y não precisa mais ser um valor regular de $F|U$.]

Sejam portanto $F\colon \overline{U} \to \mathbb{R}^m$ contínua e $y \in \mathbb{R}^m - F(\partial U)$. Para definir $gr(F, U, y)$, tomemos $\varepsilon > 0$, tal que $|y - F(x)| > 2\varepsilon$ para todo $x \in \partial U$, e uma aplicação $G\colon \overline{U} \to \mathbb{R}^m$, de classe C^∞, com $|G(x) - F(x)| < \varepsilon$ para qualquer $x \in \overline{U}$.

Afirmamos que $|z - y| < \varepsilon \Rightarrow z \notin G(\partial U)$, ou seja, que $B(y; \varepsilon) \cap G(\partial U) = \varnothing$. Com efeito,

$$[z \in B(y; \varepsilon), x \in \partial U] \Rightarrow 2\varepsilon < |y - F(x)| \leq |y - z| + |z - G(x)| +$$
$$+ |G(x) - F(x)| \Rightarrow 2\varepsilon < \varepsilon + |z - G(x)| + \varepsilon, \text{ donde } |z - G(x)| > 0.$$

Pomos então, por definição $gr(F, U, y) = gr(G, U, z)$, onde z é qualquer valor regular de G tal que $|z - y| < \varepsilon$.

Resta mostrar que esta definição é legítima. Em primeiro lugar, existem valores regulares de G na bola $B(y; \varepsilon)$, em virtude do Teorema de Sard. Em segundo lugar $gr(G, U, z)$ é o mesmo para todos esses valores regulares z pois a bola $B(y; \varepsilon)$ é um conjunto conexo, disjunto

de $G(\partial U)$ e, como sabemos o grau $gr(G, U, z)$ é constante quando z varia na mesma componente conexa. Finalmente, mostremos que a definição dada independe da aplicação G escolhida.

Com efeito, sejam $G_1, G_2 \colon \overline{U} \to \mathbb{R}^m$ aplicações C^∞ tais que $|G_1(x) - F(x)| < \varepsilon$ e $|G_2(x) - F(x)| < \varepsilon$ para todo $x \in \overline{U}$. Então

$$|(1-t)G_1(x) + tG_2(x) - F| \leq (1-t)|G_1(x) - F(x)| + t|G_2(x) - F(x)| < \varepsilon$$

para todo $t \in [0,1]$ e todo $x \in \overline{U}$. Assim, a homotopia linear $H \colon I \times \overline{U} \to \mathbb{R}^m$, $H(t,x) = (1-t)G_1(x) + t \cdot G_2(x)$, entre G_1 e G_2, é tal que $|H(t,x) - F(x)| < \varepsilon$ para todo $t \in [0,1]$ e todo $x \in \overline{U}$, logo, como vimos acima, $B(y;\varepsilon) \cap H(I \times \partial U) = \varnothing$ e daí, para todo $z \in B(y;\varepsilon)$ que seja valor regular de G_1 e G_2, tem-se $gr(G_1, U, z) = gr(G_2, U, z)$.

Assim, $gr(F, U, y)$ está bem definido, para todo $y \in \mathbb{R}^m - F(\partial U)$. Vejamos agora as propriedades do grau da aplicação contínua $F \colon \overline{U} \to \mathbb{R}^m$.

A) *$gr(F, U, y)$, como função de y, é constante em cada componente conexa de $\mathbb{R}^m - F(\partial U)$.*

Basta provar que $gr(F, U, y)$, como função de $y \in \mathbb{R}^m - F(\partial U)$, é localmente constante. Ora, se $|y - F(x)| > 2\varepsilon$ para todo $x \in \partial U$, sabemos que $gr(F, U, y) = gr(G, U, z)$, onde $G \colon \overline{U} \to \mathbb{R}^m$ é de classe C^∞, com $|G(x) - F(x)| < \varepsilon$ para todo $x \in \overline{U}$, e $z \in B(y;\varepsilon)$ é qualquer valor regular de G. Afirmamos que, para todo $w \in B(y;\varepsilon)$, tem-se $gr(F, U, w) = gr(F, U, y)$. Com efeito, vale $|w - F(x)| > \varepsilon$ para todo $x \in \partial U$. Logo, se tomarmos $G \colon \overline{U} \to \mathbb{R}^m$ de classe C^∞, tal que $|G(x) - F(x)| < \varepsilon/2$ para todo $x \in \overline{U}$, teremos $gr(F, U, w) = gr(G, U, v)$ seja qual for o valor regular $v \in B(w;\varepsilon/2)$ de G. Escolhemos então $u \in B(y;\varepsilon) \cap B(w;\varepsilon/2)$, valor regular de G e teremos $gr(F, U, y) = gr(G, U, u) = gr(F, U, w)$, assim completando a demonstração.

B) *Se $gr(F, U, y) \neq 0$ então $y \in F(U)$.*

Do contrário, seria $y \notin F(\overline{U})$, pois $y \notin F(\partial U)$. Como $F(\overline{U})$ é compacto, existiria $r > 0$ tal que $|y - F(x)| > r$ para todo $x \in \overline{U}$. Então, seja qual fosse a aplicação $G \colon \overline{U} \to \mathbb{R}^m$, de classe C^∞, com $|G(x) - F(x)| < r/2$ para todo $x \in \overline{U}$, teríamos $B(y; r/2) \cap G(\overline{U}) = \varnothing$. Assim, para todo valor regular z de G suficientemente próximo de y, teríamos $z \notin G(\overline{U})$, portanto $gr(F, U, z) = gr(G, U, z) = 0$.

C) *Se $gr(F, U, y) \neq 0$ então $y \in \operatorname{int.} F(U)$.*

Seja V a componente conexa de $\mathbb{R}^m - F(\partial U)$ que contém y. Por A), $gr(F, U, z) \neq 0$ para todo $y \in V$. Por B), $V \subset F(U)$. Como V é um aberto contendo y, temos $y \in \text{int.}\, F(U)$.

D) *Se existe uma homotopia (de classe C^0) $H: I \times \overline{U} \to \mathbb{R}^m$, entre F e G, tal que $y \notin H(I \times \partial U)$, então $gr(F, U, y) = gr(G, U, y)$ para todo $y \in \mathbb{R}^m - [F(\partial U) \cup G(\partial U)]$.*

Com efeito, seja $\varepsilon > 0$ tal que $|y - F(x)| > 2\varepsilon$ para todo $x \in \partial U$. Sejam $\overline{F}, \overline{G}: \overline{U} \to \mathbb{R}^m$ de classe C^∞, tais que $|\overline{F}(x) - F(x)| < \varepsilon$ e $|\overline{G}(x) - G(x)| < \varepsilon$ para todo $x \in \overline{U}$. Então $gr(F, U, y) = gr(\overline{F}, U, z)$ e $gr(G, U, z) = gr(\overline{G}, U, z)$, onde z é qualquer ponto de $B(y; \varepsilon)$ que seja valor regular de $\overline{F}$ e $\overline{G}$ ao mesmo tempo. Ora, as imagens de ∂U pelas homotopias lineares $\overline{F} \simeq F$ e $\overline{G} \simeq G$ não contêm y. Por transitividade, obtemos uma homotopia de classe C^0, $\overline{F} \simeq \overline{G}$, pela qual a imagem de ∂U não contém y. Por aproximação, chegamos a uma homotopia $L: I \times \overline{U} \to \mathbb{R}^m$ de classe C^∞ entre $\overline{F}$ e $\overline{G}$, tal que $L(I \times \partial U)$ não contém y. Daí resulta então que $gr(\overline{F}, U, z) = gr(\overline{G}, U, z)$.

E) *Sejam $F, G: \overline{U} \to \mathbb{R}^m$ aplicações contínuas tais que $F(x) = G(x)$ para todo $x \in \partial U$. Então $gr(F, U, y) = gr(G, U, y)$ qualquer que seja $y \in \mathbb{R}^m - F(\partial U) = \mathbb{R}^m - G(\partial U)$.*

Isto resulta de D), mediante a homotopia linear $H: I \times \overline{U} \to \mathbb{R}^m$, $H(x, t) = (1 - t)F(x) + t \cdot G(x)$.

Exemplo 36. Seja $f: U \to \mathbb{R}^m$ contínua no aberto $U \subset \mathbb{R}^m$, diferenciável no ponto $a \in U$ e tal que $f'(a): \mathbb{R}^m \to \mathbb{R}^m$ seja um isomorfismo. Existe $\delta > 0$ tal que $B = B[a; \delta] \subset U$ e $x \in B \Rightarrow f(x) - f(a) = f'(a) \cdot (x - a) + r(x)$, com $|r(x)| < |f'(a) \cdot (x - a)|$. Escrevendo $g(x) = f(x) - f(a)$, vemos que a é o único zero de g em B. Em particular $0 \notin g(\partial B)$. Além disso, a origem $0 \in \mathbb{R}^m$ não pertence à imagem da homotopia linear entre g e a transformação $T: \mathbb{R}^m \to \mathbb{R}^m$, dada por $T \cdot x = f'(a) \cdot (x - a)$. Logo $gr(g, B, 0) = gr(T, B, 0) = \pm 1$, esta última igualdade tendo sido provada na demonstração de K5). Segue-se da propriedade C), acima, que $0 \in \text{int.}\, g(B)$. Daí, $a \in \text{int.}\, f(B)$ e portanto $a \in \text{int.}\, f(U)$. Isto mostra a necessidade de f ser descontínua no Exemplo 9 do Capítulo V.

APÊNDICE

Seja $G = Gl(\mathbb{R}^m)$ o conjunto das matrizes invertíveis $m \times m$. Os conjuntos $G_+ = \{X \in G; \det. X > 0\}$ e $G_- = \{X \in G; \det. X < 0\}$ são abertos disjuntos. Como $G = G_+ \cup G_-$, vemos que G é desconexo.

Mostraremos agora que G_+ e G_- são conexos por caminhos, logo são as componentes conexas de G. Se A é qualquer matriz $m \times m$ com determinante negativo, a aplicação $X \mapsto A \cdot X$ é um homeomorfismo entre G_+ e G_- . Assim, basta provar que G_+ é conexo por caminhos.

Para o conjunto $O(m)$ das matrizes ortogonais $m \times m$, temos uma decomposição análoga $O(m) = O_+(m) \cup O_-(m)$ e, de modo semelhante (tomando A ortogonal acima), vemos que $O_+(m)$ e $O_-(m)$ são homeomorfos. Provaremos primeiro que $O_+(m)$ é conexo por caminhos. Esta é a segunda das proposições abaixo. Na primeira, mostramos que, em qualquer superfície M, se pode transladar uma base ortonormal de T_pM ao longo de um caminho em M começando no ponto p.

Proposição A. *Dados um caminho $\lambda\colon [0,1] \to M$ numa superfície $M \subset \mathbb{R}^n$ e uma base ortonormal $\{e_1, \ldots, e_m\} \subset T_{\lambda(0)}M$, existem m caminhos $v_1, \ldots, v_m\colon [0,1] \to \mathbb{R}^n$ tais que $v_1(0) = e_1, \ldots, v_m(0) = e_m$ e, para cada $t \in [0,1]$, $\{v_1(t), \ldots, v_m(t)\}$ é uma base ortonormal em $T_{\lambda(t)}M$.*

Demonstração: Suponhamos, inicialmente, que exista uma parametrização $\varphi\colon U_0 \to U$ em M tal que $\lambda(t) \in U$, para todo $t \in [0,1]$. Assim, faz sentido definir o caminho $\lambda_0 = \varphi^{-1} \circ \lambda\colon [0,1] \to U_0$. Seja $\{u_1, \ldots, u_m\} \subset \mathbb{R}^m$ uma base tal que $\varphi'(\lambda_0(0)) \cdot u_i = e_i$ $(i = 1, \ldots, m)$. Definamos os caminhos $w_1, \ldots, w_m\colon [0,1] \to \mathbb{R}^n$ pondo $w_i(t) = \varphi'(\lambda_0(t)) \cdot u_i$. Para cada $t \in [0,1]$, $\{w_1(t), \ldots, w_m(t)\}$ é uma base de $T_{\lambda(t)}M$, à qual aplicamos o processo de ortonormalização de Gram-Schmidt para obter a base $\{v_1(t), \ldots, v_m(t)\}$ procurada. O caso geral se reduz a este mediante a observação de que, dado um caminho $\lambda\colon [0,1] \to M$, existe uma partição $0 = t_0 < t_1 < \cdots < t_k = 1$ tal que, para cada $i = 1, 2, \ldots, k$, a imagem $\lambda([t_{i-1}, t_i])$ está contida numa vizinhança parametrizada.

Para demonstrar a proposição seguinte, usaremos a Proposição A no caso particular em que M é a esfera S^{m-1}, observando que $\{w_1, \ldots, w_m\}$ é uma base ortonormal em $\mathbb{R}^m$ se, e somente, se $\{w_1, \ldots, w_{m-1}\}$ é uma base ortonormal em $T_{w_m}(S^{m-1})$.

Proposição B. *Sejam $\{u_1, \ldots, u_m\}$ e $\{v_1, \ldots, v_m\}$ bases ortonormais positivas em $\mathbb{R}^m$. Existem caminhos $w_1, \ldots, w_m\colon [0,2] \to \mathbb{R}^m$ tais que $w_i(0) = u_i$, $w_i(2) = v_i$ e, para todo $t \in [0,2]$, $\{w_1(t), \ldots, w_m(t)\}$ é uma*

base ortonormal positiva de $\mathbb{R}^m$*. Como uma matriz* $m \times m$ *é ortogonal se, e somente se, seus vetores-coluna formam uma base ortonormal de* $\mathbb{R}^m$*, isto equivale a afirmar que* $O_+(m)$ *é conexo por caminhos.*

Demonstração: (Indução em m.) O fato é óbvio se $m = 1$. Supondo-o válido para $m-1$, tomemos um caminho $\lambda \colon [0,1] \to S^{m-1}$, com $\lambda(0) = u_m$ e $\lambda(1) = v_m$. Sendo $\{u_1, \ldots, u_{m-1}\}$ uma base ortonormal em $T_{u_m}(S^{m-1})$, a Proposição A nos dá os caminhos $w_1, \ldots, w_{m-1} \colon [0,1] \to \mathbb{R}^m$, tais que $w_1(0) = u_1, \ldots, w_{m-1}(0) = u_{m-1}$ e, para todo $t \in [0,1]$, $\{w_1(t), \ldots, w_{m-1}(t)\}$ é uma base ortonormal de $T_{\lambda(t)}(S^{m-1})$, logo $\{w_1(t), \ldots, w_{m-1}(t), \lambda(t)\}$ é uma base ortonormal de $\mathbb{R}^m$ (necessariamente positiva, pois o determinante que tem esses m vetores por colunas é uma função contínua não-nula de t, positiva para $t = 0$). A matriz de passagem da base $\{w_1(1), \ldots, w_{m-1}(1), \lambda(1)\}$ para $\{v_1, \ldots, v_m\}$ é ortogonal, com determinante positivo e, como $\lambda(1) = v_m$, tem a forma $\widetilde{A} = \begin{pmatrix} 1 & 0 \\ 0 & A \end{pmatrix}$ onde A é a matriz de passagem de $\{w_1(1), \ldots, w_{m-1}(1)\}$ para $\{v_1, \ldots, v_{m-1}\}$, bases ortonormais de $T_{v_m}(S^{m-1})$. Como det. $A =$ det. $\widetilde{A} > 0$, segue-se que $A \in O_+(m-1)$. Pela hipótese de indução, existe um caminho de matrizes $A(t) = (a_{ij}(t)) \in O_+(m-1)$, $1 \le t \le 2$, tal que $A(1) =$ matriz identidade e $A(2) = A$. Pondo $w_i(t) = \sum\limits_{j=1}^{m-1} a_{ij}(t) w_j(1)$, obtemos $m-1$ caminhos $w_1, \ldots, w_{m-1} \colon [1,2] \to \mathbb{R}^m$, com $w_1(2) = v_1, \ldots, w_{m-1}(2) = v_{m-1}$. Escrevendo $w_m(t) = v_m \quad (1 \le t \le 2)$ completamos a definição dos caminhos $w_1, \ldots, w_m \colon [0,2] \to \mathbb{R}^m$ que procurávamos.

Proposição C. *Sejam* $\{a_1, \ldots, a_m\}$ *e* $\{b_1, \ldots, b_m\}$ *bases positivas de* $\mathbb{R}^m$*. Existem* $c_1, \ldots, c_m \colon [0,1] \to \mathbb{R}^m$ *tais que* $c_i(0) = a_i$, $c_i(1) = b_i$ *e, para cada* $t \in [0,1]$, $\{c_1(t), \ldots, c_m(t)\} \subset \mathbb{R}^m$ *é uma base positiva. Associando-se a cada matriz* $m \times m$ *a lista dos seus vetores-coluna, isto equivale a dizer que o conjunto* G_+ *das matrizes* $m \times m$ *com determinante positivo é conexo por caminhos.*

Demonstração: Mostraremos inicialmente que, dada qualquer base $\{a_1, \ldots, a_m\} \subset \mathbb{R}^m$, existem caminhos $d_1, \ldots, d_m \colon [0,2] \to \mathbb{R}^m$ tais que, para cada $t \in [0,2]$, $\{d_1(t), \ldots, d_m(t)\} \subset \mathbb{R}^m$ é uma base, com $d_i(0) = a_i$ e $\{d_1(2), \ldots, d_m(2)\} \subset \mathbb{R}^m$ ortonormal. Isto significa que toda matriz invertível A pode ser ligada a uma matriz ortogonal por um caminho de matrizes invertíveis $A(t)$, $0 \le t \le 2$. [Evidentemente, det. $A(t)$ terá

o mesmo sinal, para todo t.] A prova será por indução em m. O caso $m = 1$ é trivial. Supondo o resultado verdadeiro para $\mathbb{R}^{m-1}$, ele valerá também para qualquer subespaço $(m-1)$-dimensional de $\mathbb{R}^m$. Seja então E o subespaço de $\mathbb{R}^m$ gerado por $a_1, \ldots, a_{m-1}$. Consideremos um vetor unitário $e \in \mathbb{R}^m$ perpendicular a E. Substituindo e por $-e$, se necessário, podemos supor que $e = \alpha_1 a_1 + \cdots + \alpha_m a_m$, com $\alpha_m > 0$. Para todo $t \in [0,1]$, ponhamos $d_i(t) = a_i$ se $1 \leq i \leq m-1$ e $d_m(t) = (1-t)a_m + t \cdot e$. Para todo $t \in [0,1]$, o coeficiente de a_m em $d_m(t)$ é $1 - t + t\alpha_m > 0$, logo $d_m(t) \notin E$ e assim $\{d_1(t), \ldots, d_m(t)\} \subset \mathbb{R}^m$ é uma base, igual a $\{a_1, \ldots, a_m\}$ para $t = 0$ e igual a $\{a_1, \ldots, a_{m-1}, e\}$ para $t = 1$. Pela hipótese de indução, existem caminhos $d_1, \ldots, d_{m-1} \colon [1,2] \to E$ tais que, para todo $t \in [1,2]$, $\{d_1(t), \ldots, d_{m-1}(t)\} \subset E$ é uma base, igual a $\{a_1, \ldots, a_{m-1}\}$ quando $t = 1$ e ortonormal quando $t = 2$. Então, pondo $d_m(t) = e$ para todo $t \in [1,2]$, completamos a definição dos caminhos $d_1, \ldots, d_m \colon [0,2] \to \mathbb{R}^m$, da forma como queríamos. A demonstração se completa invocando a Proposição B: deformamos as bases dadas em bases ortonormais, como acabamos de provar e, usando a Proposição B, ligamos estas bases ortonormais entre si. (Evidentemente, se (a_i) e (b_j) são bases positivas, também serão positivas as bases ortonormais que delas se obtêm por deformação.)

Exercícios

§1.

1.1. Se $\varphi \colon E \times \cdots \times E \to F$ é r-linear então $\varphi(v_1, \ldots, v_r) = 0$ quando algum dos v_i é zero.

1.2. Seja $\mathcal{L}_r(E;F)$ o espaço vetorial das aplicações r-lineares de E em F. Dada uma base de E, obtenha a partir dela uma base para $\mathcal{L}_r(E;\mathbb{R})$ e, em seguida, para $\mathcal{L}_r(E;\mathbb{R}^n)$. Determine a dimensão de $\mathcal{L}_r(E;F)$ em função das dimensões de E e F.

1.3. Seja $\varphi \colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^4$ a aplicação bilinear definida por $\varphi(x_1, x_2, y_1, y_2) = (x_1y_1, x_1y_2, x_2y_1, x_2y_2)$. Mostre que a imagem de φ não é um subespaço vetorial.

1.4. Se $\{e_1, e_2, e_3, e_4\}$ é uma base de $(\mathbb{R}^4)^*$, prove que $e_1 \wedge e_2 + e_3 \wedge e_4$ não é decomponível.

1.5. As seguintes afirmações a respeito de uma aplicação r-linear $\varphi \colon E \times \cdots \times E \to F$ são equivalentes:

(a) φ é alternada;

(b) $\varphi(\dots, v_i, v_{i+1}, \dots) = -\varphi(\dots, v_{i+1}, v_i, \dots)$;

(c) $\varphi(\dots, v, v \dots) = 0$.

1.6. A aplicação r-linear $\varphi\colon E \times \cdots \times E \to F$ é alternada se, e somente se, para toda permutação σ do conjunto dos inteiros $\{1, \dots, r\}$, e toda lista de vetores $v_1, \dots, v_r \in E$, vale $\varphi(v_{\sigma(1)}, \dots, v_{\sigma(r)}) = \varepsilon_\sigma \cdot \varphi(v_1, \dots, v_r)$, onde $\varepsilon_\sigma = 1$ se σ é uma permutação par e $\varepsilon_\sigma = -1$ se a permutação σ é ímpar.

1.7. Exprimir a 2-forma $e_1 \wedge e_2 + e_3 \wedge e_4$ de várias maneiras distintas como soma de formas decomponíveis. (Aqui $\{e_1, e_2, e_3, e_4\}$ é uma base do dual de $\mathbb{R}^4$.)

1.8. O posto da transformação linear $A\colon E \to F$ é o maior inteiro r tal que $A^*\colon \mathcal{A}_r(F) \to \mathcal{A}_r(E)$ é diferente de zero.

1.9. Seja $\varphi\colon \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ uma forma alternada de grau 2. Se m é ímpar, existe um vetor não-nulo $v \in \mathbb{R}^m$ tal que $\varphi(v, w) = 0$ para todo $w \in \mathbb{R}^m$.

1.10. Dados arbitrariamente $a_1, \dots, a_m$ reais, com $a_1 \neq 0$, e uma base $\{e_1, \dots, e_m\} \subset E^*$, ponha $f_1 = a_2 e_1 + a_1 e_2$ e $f_i = (-1)^{i+1}(a_{i+1}/a_1)e_1 + e_{i+1}$, $2 \leq i \leq m-1$. Mostre que

$$f_1 \wedge \cdots \wedge f_{m-1} = \sum_{i=1}^{m} a_i e_1 \wedge \cdots \wedge e_{i-1} \wedge e_{i+1} \wedge \cdots \wedge e_m\,.$$

Conclua que se $\dim E = m$ então toda $\omega \in \mathcal{A}_{m-1}(E)$ é decomponível.

1.11. Seja E um espaço vetorial de dimensão m, orientado e munido de um produto interno. Dado $v \in E$, seja $\omega = \varphi(v) \in \mathcal{A}_{m-1}(E)$ definida por

$$\omega(v_1, \dots, v_{m-1}) = \langle v, v_1 \times \cdots \times v_{m-1} \rangle.$$

Mostre que a aplicação $\varphi\colon E \to \mathcal{A}_{m-1}(E)$, assim obtida, é um isomorfismo. Se $v \in E$ é o primeiro elemento de uma base ortonormal positiva cuja dual é $\{e_1, \dots, e_m\}$ então $\varphi(v) = e_2 \wedge \cdots \wedge e_m$. Conclua de novo que se $\dim E = m$ então toda $\omega \in \mathcal{A}_{m-1}(E)$ é decomponível.

§2.

2.1. Fixados, $k, m \in \mathbb{N}$, $k \geq r$, seja ω uma forma de grau r em $\mathbb{R}^m$, tal que $f^*\omega = 0$ para toda transformação afim $f\colon \mathbb{R}^k \to \mathbb{R}^m$. Prove que $\omega = 0$.

2.2. Sejam a e b matrizes $r \times m$ e $m \times r$ respectivamene, com $r \leq m$. Prove que $\det(ab) = \sum_J \det a_J \cdot \det b_j$, onde J percorre os subconjuntos com r elementos do conjunto $\{1, 2, \dots, m\}$. Conseqüência: se $u_1, \dots, u_m$, $v_1, \dots, v_m \in \mathbb{R}^{m+1}$ então

$$\det(\langle u_i, v_j \rangle) = \langle u_1 \times \cdots \times u_m,\, v_1 \times \cdots \times v_m \rangle.$$

2.3. Interprete a identidade de Lagrange como uma generalização do Teorema de Pitágoras: o quadrado do volume de um paralelepípedo r-dimensional em $\mathbb{R}^m$ é igual à soma dos quadrados dos volumes das suas $\binom{m}{r}$ projeções sobre os subespaços de $\mathbb{R}^m$ gerados por r vetores da base canônica.

2.4. Sejam ω uma forma de grau 1 contínua no aberto $U \subset \mathbb{R}^m$ e $\lambda\colon [a,b] \to U$ um caminho de classe C^1. Tem-se $\lambda^*\omega = f dx$, onde $f\colon [a,b] \to \mathbb{R}$ é contínua e $\{dx\}$ é a base canônica de $\mathbb{R}^*$. Mostre que a integral curvilínea $\int_\lambda \omega$, definida no Capítulo IV, nada mais é do que $\int_a^b f(x)dx$.

2.5. Sejam α, β formas de grau 1 num aberto $U \subset \mathbb{R}^3$ tais que $\alpha \wedge \beta \neq 0$ em todos os pontos de U. Se uma forma ω, de grau 2 em U é tal que $\omega \wedge \alpha = \omega \wedge \beta = 0$, prove que existe uma função $f\colon U \to \mathbb{R}$ tal que $\omega = f \cdot \alpha \wedge \beta$. Se $\alpha, \beta, \omega \in C^1$ então $f \in C^1$.

2.6. Sejam α, ω formas de classe C^1 e de grau 1 em $\mathbb{R}^3$. Se $\omega(x) \neq 0$ para todo $x \in \mathbb{R}^3$ e $\alpha \wedge \omega = 0$ então $\alpha = f\omega$ onde $f\colon \mathbb{R}^3 \to \mathbb{R}$ é uma função de classe C^1.

§3.

3.1. Seja ω uma forma diferencial C^∞ de grau r num aberto U do espaço euclidiano. Se ω se anula sobre os vetores tangentes a uma superfície contida em U, o mesmo ocorre com $d\omega$.

3.2. Para toda r-forma ω de classe C^k em S^m existe uma r-forma Ω de classe C^k em $\mathbb{R}^{m+1} - \{0\}$ tal que $\Omega|S^m = \omega$. Se $m > 1$, conclua que toda forma diferencial fechada de grau 1 na esfera S^m é a diferencial de uma função $f\colon S^m \to \mathbb{R}$. Em particular, toda forma fechada de grau 1 em S^m $(m > 1)$ deve anular-se em pelo menos dois pontos. Usando uma vizinhança tubular, estenda os dois últimos resultados para toda superfície compacta simplesmente conexa.

3.3. Sejam ω uma 1-forma de classe C^∞ em $\mathbb{R}^m$ e $f\colon \mathbb{R}^m \to \mathbb{R}$ uma função C^∞, com $f(x) \neq 0$ para todo $x \in \mathbb{R}^m$. Mostre que $d(f\omega) \equiv 0$ se, e somente se, a 1-forma $\alpha = \omega - (1/f)dx_{m+1}$ em $\mathbb{R}^{m+1}$ cumpre a condição $\alpha \wedge d\alpha = \omega \wedge d\omega$. (Considere $\mathbb{R}^m \subset \mathbb{R}^{m+1}$ definido por $x_{m+1} = 0$.)

3.4. Seja $U \subset \mathbb{R}^m$ aberto. Represente um ponto de $U \times \mathbb{R}$ pela notação $(x,t) = (x_1, \ldots, x_m, t)$. Toda forma ω, de grau r em $U \times \mathbb{R}$, escreve-se de modo único como $\omega = dt \wedge \alpha + \beta$, onde $\alpha(x,t) = \sum\limits_I a_I(x,t)dx_I$ e $\beta(x,t) = \sum\limits_J b_J(x,t)dx_J$ são formas de grau $r-1$ e r, respectivamente. Defina uma transformação linear K, que associa a cada forma contínua de grau r em $U \times \mathbb{R}$ uma forma contínua de grau $r-1$ em U, pondo $(K\omega)(x) = \int_0^1 \alpha(x,t)dt = \sum\limits_I (\int_0^1 a_I(x,t)dt)dx_I$. Para cada $t \in \mathbb{R}$, considere a aplicação $i_t\colon U \to U \times \mathbb{R}$, definida por $i_t(x) = (x,t)$. Prove que, para cada forma ω, de classe C^1 em $U \times \mathbb{R}$ tem $K(d\omega) + d(K\omega) = i_1^*(\omega) - i_0^*(\omega)$. Conclua que, a cada homotopia $H\colon U \times [0,1] \to V \subset \mathbb{R}^n$, de classe C^∞, entre as aplicações $f, g\colon U \to V$, se pode associar uma transformação linear $L = K \circ H^*$, que leva formas de grau r em V sobre formas de grau $r-1$ em U, de tal modo que $L(d\omega) + d(L\omega) = g^*\omega - f^*\omega$, para toda forma ω, de classe C^∞ em V. Daí resulta que, se ω é fechada então a diferença $g^*\omega - f^*\omega$ é exata. Em particular se o aberto U é C^∞-contrátil, então toda forma fechada de grau > 0 em U é exata. Mais particularmente, isto ocorre com um aberto convexo U e, em especial, com o espaço $\mathbb{R}^m$. ("Lema de Poincaré".) (Obs. Estenda H a uma aplicação $\overline{H}\colon U \times \mathbb{R} \to V$ pondo $\overline{H}(x,t) = H(x, \xi(t))$, com $0 \leq \xi(t) \leq 1$, $\xi(t) = 0$ se $t \leq 0$ e $\xi(t) = 1$ se $t \geq 1$.)

3.5. No exercício acima, substitua U por uma superfície M e mostre que o argumento se aplica *ipsis literis*. Mostre ainda que se as aplicações $f, g\colon M \to N$ e a homotopia $H\colon M \times I \to N$ são próprias e ω tem suporte compacto, então $f^*\omega$, $g^*\omega$, $H^*\omega$, $K\omega$ e $d\omega$ têm suporte compacto.

3.6. Seja M uma superfície C^∞. Duas formas α, β fechadas, de grau r e classe C^∞ em M dizem-se *cohomólogas* quando $\alpha - \beta$ é exata. Escreve-se então $\alpha \sim \beta$. Mostre que esta é uma relação de equivalência e que o conjunto das classes de equivalência $[\alpha]$ de formas de grau r e classe C^∞ na superfície M constitui um espaço vetorial, indicado com a notação $H^r(M)$ e chamado o *grupo de cohomologia* de grau r da superfície M. Mostre que toda aplicação $f\colon M \to N$ de classe C^∞ entre superfícies induz, para cada $r \geq 0$ inteiro, uma transformação linear $f^*\colon H^r(N) \to H^r(M)$, tal que $(g \circ f)^* = f^* \circ g^*$, definida por $f^*[\alpha] = [f^*\alpha]$. Se f e g são homotópicas em classe C^2, então $f^* = g^*$. Mostre que $H^r(M) = \{0\}$ se $r >$ dim. M. Mostre ainda que $\alpha \sim \beta$, $\alpha' \sim \beta' \Rightarrow \alpha \wedge \alpha' \sim \beta \wedge \beta'$, logo o produto exterior de formas induz uma aplicação bilinear $([\alpha], [\beta]) \mapsto [\alpha \wedge \beta]$ de $H^r(M) \times H^s(M)$ em $H^{r+s}(M)$, a qual torna a soma direta $H^*(M) = H^0(M) \oplus \cdots \oplus H^m(M)$ $(m = \dim M)$ uma álgebra associativa e anti-comutativa. Mostre que a transformação linear $f^*\colon H^*(N) \to H^*(M)$, induzida por uma aplicação $f\colon M \to N$ de classe C^∞, é um homomorfismo de álgebras. Mostre, finalmente, que se M é conexa então $H^0(M)$ tem dimensão 1.

3.7. Seja M uma superfície de classe C^∞. Considerando exclusivamente formas de classe C^∞ com suporte compacto em M, defina os *grupos de cohomologia com suporte compacto* $H_c^k(M)$ e estabeleça para os mesmos propriedades análogas às que foram enunciadas no exercício anterior, supondo porém que as aplicações $f\colon M \to N$ e as homotopias $H\colon M \times I \to N$ em questão são próprias.

3.8. Seja p um ponto arbitrariamente fixado numa superfície M, de classe C^∞. Mostre que a aplicação identidade $i\colon M \to M$ é C^∞-homotópica a uma aplicação $f\colon M \to M$ que é constante, igual a p, em todos os pontos de uma vizinhança de p e conclua que toda forma fechada α de classe C^∞ em M é cohomóloga a uma forma β, também de classe C^∞, tal que $p \notin$ supp. β. Conclua que $H^k(S^m)$ é isomorfo a $H_c^k(\mathbb{R}^m)$ para todo $k = 1, 2, \ldots$.

3.9. Diz-se que duas superfícies M, N, de classe C^∞, têm o mesmo *tipo de homotopia* C^∞ quando existem aplicações $f\colon M \to N$ e $g\colon N \to M$, de classe C^∞, tais que $g \circ f\colon M \to M$ e $f \circ g\colon N \to N$ são C^∞-homotópicas às aplicações identidade correspondentes. Neste caso, f chama-se uma *equivalência homotópica* e g a sua inversa homotópica. Prove que, neste caso, $f^*\colon H^k(N) \to H^k(M)$ e $g^*\colon H^k(M) \to H^k(N)$ são, para $k = 0, 1, \ldots$, isomorfismos, um inverso do outro. Mostre que a inclusão $i\colon S^m \to \mathbb{R}^{m+1} - \{0\}$ e a projeção radial $\rho\colon \mathbb{R}^{m+1} - \{0\} \to S^m$ (definida por $\rho(x) = x/|x|$) são equivalências homotópicas entre S^m e $\mathbb{R}^{m+1} - \{0\}$, uma inversa da outra. Conclua que os grupos de cohomologia $H^k(S^m)$ e $H^k(\mathbb{R}^{m+1} - \{0\})$ são isomorfos, para $k = 0, 1, \ldots$.

3.10. Uma função contínua $f\colon \mathbb{R} \to \mathbb{R}$, de suporte compacto, é a derivada de uma função com suporte compacto se, e somente se, $\int_{\mathbb{R}} f(x)dx = 0$. Conclua que a aplicação $f \mapsto \int_{\mathbb{R}} f$ induz um isomorfismo entre $H_c^1(\mathbb{R})$ e $\mathbb{R}$. Daí deduza um isomorfismo entre $H^1(S^1)$ e $\mathbb{R}$ a partir da aplicação $\omega \mapsto \int_{S_1} \omega$. Finalmente,

usando a equivalência homotópica entre S^1 e $\mathbb{R}^2 - \{0\}$, reobtenha, de modo diferente, o resultado do §10, Capítulo IV, segundo o qual $H^1(\mathbb{R}^2 - \{0\})$ é isomorfo a $\mathbb{R}$ e, mais precisamente, uma 1-forma fechada ω em $\mathbb{R}^2 - \{0\}$ é exata se, e somente se $\int_\lambda \omega = 0$, onde λ é um caminho seccionalmente C^1 em $\mathbb{R}^2 - \{0\}$, tal que $n(\lambda; 0) \neq 0$.

§4.

4.1. Seja $(C_\lambda)_{\lambda \in L}$ uma família de subconjuntos de uma superfície M tal que cada compacto $K \subset M$ tem interseção não-vazia com apenas um número finito de conjuntos C_λ. Prove que a família dada é localmente finita.

4.2. Seja $(F_\lambda)_{\lambda \in L}$ uma família localmente finita de subconjuntos fechados de uma superfície M. Ponha $F = \cup F_\lambda$. Se uma aplicação $f\colon F \to \mathbb{R}^n$ é tal que $f|F_\lambda$ é contínua para cada λ então f é contínua.

4.3. Sejam F fechado, U aberto, com $F \subset U \subset \mathbb{R}^m$. Prove que existe uma função $f\colon \mathbb{R}^m \to [0,1]$ de classe C^∞, tal que $f(x) = 0$ para todo $x \in F$ e $f(x) = 1$ para todo $x \in \mathbb{R}^m - U$.

4.4. Dado $K \subset \mathbb{R}^m$ compacto, prove que existe $f\colon \mathbb{R}^m \to \mathbb{R}$ de classe C^∞ tal que $f(x) = 0$ se, e somente se, $x \in K$. [*Sugestão*: Ache abertos $U_1 \supset U_2 \supset \dots$ tais que $K = \cap U_i$. Para cada i, seja $f_i\colon \mathbb{R}^m \to [0,1]$ de classe C^∞, nula em K e igual a 1 fora de U_i. Tome $c_i = \frac{1}{2^i(||f_i'||+1)(||f_i''||+1)\dots(||f_i^{(i)}||+1)}$ e ponha $f = \sum_i c_i f_i$.]

4.5. No exercício anterior, mostre que, dado previamente um aberto U em $\mathbb{R}^m$, com $K \subset U$, pode-se tomar a função f tal que $f(\mathbb{R}^m - U) = 1$.

4.6. Para todo subconjunto fechado $F \subset \mathbb{R}^m$, existe uma função $f\colon \mathbb{R}^m \to \mathbb{R}$, de classe C^∞, tal que $f^{-1}(0) = F$. [*Sugestão*: Para $i = 0, 1, 2 \dots$ sejam $K_i = \{x \in F; i \leq |x| \leq i+1\}$, $U_i = \{x \in \mathbb{R}^m; i-1 < |x| < i+1\}$ e $f_i\colon \mathbb{R}^m \to \mathbb{R}$ uma função C^∞ tal que $f_i^{-1}(0) = K_i$ e $f_i(\mathbb{R}^m - U_i) = 1$. Defina $f\colon \mathbb{R}^m \to \mathbb{R}$ pondo, para cada $x \in \mathbb{R}^m$, $f(x) = f_0(x) \cdot f_1(x) \cdot \dots \cdot f_i(x) \cdot \dots$. Note que este produto é localmente finito, logo $f \in C^\infty$. Observe também que $f^{-1}(0) = \bigcup_i f_i^{-1}(0) = \cup K_i = F$.]

4.7. Seja $M = U \cup V$ uma superfície de classe C^∞. Suponha dada uma r-forma ω, de classe C^∞ na interseção $U \cap V$ dos dois abertos $U, V \subset M$. Mediante uma partição da unidade, mostre que existem r-formas α em U e β em V, de classe C^∞, tais que $\alpha - \beta = \omega$ em $U \cap V$. Se ω é fechada, mostre que $d\alpha$ e $d\beta$ são respectivamente as restrições a U e a V de uma $(r+1)$-forma γ de classe C^∞ em M.

§5.

5.1. Seja $f\colon X \to \mathbb{R}^n$ uma aplicação de classe C^k $(k \geq 1)$ num subconjunto fechado $X \subset \mathbb{R}^m$. Prove que existe $F\colon \mathbb{R}^m \to \mathbb{R}^n$ de classe C^k tal que $F|X = f$. [Este fato também é verdadeiro para $k = 0$ e é conhecido como o "Teorema de Tietze".]

5.2. Sejam $B = B[0,1] \subset \mathbb{R}^m$ e $f\colon B \to \mathbb{R}^m$ uma aplicação contínua tal que $f(S^{m-1}) \subset B$. Prove que existe $x \in B$ tal que $f(x) = x$.

5.3. Seja $U = \{x \in \mathbb{R}^m; |x| < 1\}$. Dê exemplo de uma aplicação $f\colon U \to U$ de classe C^∞, tal que $f(x) \neq x$ para todo $x \in U$.

§6.

6.1. Seja $f\colon \mathbb{R}^4 \to \mathbb{R}^4$ um difeomorfismo de classe C^1 tal que $f^*(dx_1 \wedge dx_2) = 2dx_3 \wedge dx_4$ e $f^*(dx_3 \wedge dx_4) = 1/2dx_1 \wedge dx_2$. Prove que para todo conjunto J-mensurável $X \subset \mathbb{R}^4$ tem-se vol. $f(X) =$ vol. X.

6.2. Sejam α e β duas r-formas contínuas num aberto $U \subset \mathbb{R}^m$. Se $\int_M \alpha = \int_M \beta$ para toda superfície $M \subset U$ de dimensão r, compacta, com bordo, então $\alpha = \beta$.

6.3. Seja ω uma forma contínua de grau r num aberto $U \subset \mathbb{R}^m$. Diz-se que ω é *regular* quando existe em U uma forma contínua $\bar{\omega}$, de grau $r+1$, tal que $\int_M \bar{\omega} = \int_{\partial M} \omega$ para toda superfície M, de classe C^2, com bordo, contida em U. (1) Prove que se ω é regular então $\bar{\omega}$ é univocamente determinada por ω. (2) Prove que toda forma regular de grau zero (função) é diferenciável. (3) Sejam $f, g\colon \mathbb{R} \to \mathbb{R}$ funções contínuas quaisquer. A forma $\omega(x,y) = f(x)dx + g(y)dy$ é regular em $\mathbb{R}^2$ mas pode não ser diferenciável. (4) Se ω é regular então $\bar{\omega}$ também é, com $\bar{\bar{\omega}} = 0$. (5) A soma de duas formas regulares é regular, e vale $\overline{\omega + \varphi} = \bar{\omega} + \bar{\varphi}$. (6) Se ω é de classe C^1 então ω é regular e $\bar{\omega} = d\omega$.

6.4. Uma forma contínua ω de grau r no aberto $U \subset \mathbb{R}^m$ é regular se, e somente se, para todo aberto V, com fecho compacto $\overline{V} \subset U$, existe uma seqüência de formas ω_k de grau r, com coeficientes diferenciáveis, tais que $\omega_k \to \omega$ e $d\omega_k \to \bar{\omega}$ uniformemente em V (isto é, cada coeficiente converge uniformemente em V). Conclua que o produto exterior de duas formas regulares é regular e $\overline{\omega_1 \wedge \omega_2} = \bar{\omega}_1 \wedge \omega_2 + (-1)^{r_1} \omega_1 \wedge \bar{\omega}_2$.

6.5. Sejam U, V abertos em $\mathbb{R}^m$, $f\colon U \to V$ uma aplicação diferenciável e ω uma forma regular em V. Então $f^*\omega$ é regular e $\overline{f^*\omega} = f^*\bar{\omega}$.

6.6. Seja $M \subset \mathbb{R}^3$ uma hiperfície compacta de classe C^2. Prove que, para todo $x \in M$, a derivada da aplicação normal de Gauss é uma transformação linear auto-adjunta $\gamma'(x)\colon T_xM \to T_xM$. Prove ainda que se M é convexa (isto é, situada do mesmo lado de qualquer um dos seus planos tangentes) então, para todo $x \in M$, os autovalores de $\gamma'(x)$ são ≥ 0. Conclua que a curvatura integral de uma superfície compacta convexa é um número positivo. Generalize para $\mathbb{R}^{m+1}$ em vez de $\mathbb{R}^3$. (A única diferença essencial é distinguir m par de m ímpar.)

6.7. Sejam M uma superfície orientada de classe C^1, $f\colon M \to \mathbb{R}$ uma função contínua com suporte compacto e ω a forma elemento de volume de M. Prove que $\int_M f \cdot \omega$ é o limite das "somas de Riemann" $\Sigma f(\xi_i) \operatorname{vol}(K_i)$ onde $\xi_i \in K_i$ para cada i, $M = \cup K_i$ é uma decomposição de M como reunião finita de compactos J-mensuráveis sem pontos interiores em comum, e o limite é tomado quando a norma da decomposição (maior diâmetro dos K_i) tende a zero.

6.8. Considere o cilindro $C = \{(x,y,z) \in \mathbb{R}^3; x^2+y^2 = 1,\ 0 \leq z \leq 1\}$. Dados $m, n \in \mathbb{N}$, corte o cilindro C por $m+1$ planos equidistantes, dados pelas equações

$z = i/m$ $(i = 0, 1, \ldots, m)$, obtendo $m+1$ círculos horizontais. Subdivida cada um desses círculos em n partes iguais, de modo que os pontos de divisão de cada círculo se situem sobre os pontos médios dos arcos em que foi subdividido o círculo anterior. Inscreva no cilindro o poliedro cujos vértices são os pontos de subdivisão dos círculos, e cujas faces são os triângulos isósceles que têm por base uma corda horizontal na subdivisão feita e os lados iguais são os segmentos de reta que ligam as extremidades dessa corda ao ponto de subdivisão situado no círculo precedente ou seguinte na mesma vertical que o ponto médio do arco. Tal poliedro tem $2mn$ faces congruentes; a base de cada uma mede $2\,\mathrm{sen}(\pi/n)$ e a altura mede $\sqrt{1/m^2 + 4\,\mathrm{sen}^4\,\pi/2n}$. Mostre que $\lim\limits_{n\to\infty} A_{n,n} = 2\pi$, que $\lim\limits_{n\to\infty} A_{n^2,n} = 2\pi\sqrt{1+\pi^4/4}$ e que, para todo n fixo, $\lim\limits_{m\to\infty} A_{m,n} = \infty$. Logo não existe $\lim\limits_{m,n\to\infty} A_{m,n}$. ("Exemplo de Schwarz".)

6.9. Sejam M uma superfície orientada não-compacta de dimensão m e ω uma m-forma não-negativa, contínua em M. Uma *exaustão* de M é uma seqüência de compactos J-mensuráveis K_i tais que $M = \cup K_i$ e $K_i \subset \mathrm{int.}\,K_{i+1}$ para $i = 1, 2, \ldots$. Prove que ω é integrável se, e somente se, para toda exaustão de M existe $\lim\limits_{i\to\infty} \int_{K_i} \omega$.

6.10. Seja ω uma $(m+n)$-forma contínua com suporte compacto na superfície orientada $M \times N$, onde $\dim M = m$ e $\dim N = n$. Para cada $x \in M$ e cada lista de vetores $u_1, \ldots, u_m \in T_xM$, defina a n-forma $\beta = \omega_x(u_1, \ldots, u_m)$ em N pondo $\beta(y)(v_1, \ldots, v_n) = \omega(x,y)(u_1, \ldots, u_m, v_1, \ \ldots, v_n)$, para $y \in N$ e $v_1, \ldots, v_n \in T_xN$ quaisquer. Mostre que β é contínua, com suporte compacto, o mesmo ocorrendo com a m-forma α em M, definida por

$$\alpha(x)(u_1, \ldots, u_m) = \int_N \omega_x(u_1, \ldots, u_m) = \int_N \beta.$$

Finalmente, prove que

$$\int_{M\times N} \omega = \int_M \alpha = \int_M \left(\int_N \beta\right).$$

6.11. Sejam α e β formas contínuas com suportes compactos e graus máximos nas superfícies orientadas M e N respectivamente. Considere as projeções $\pi_M \colon M \times N \to M$ e $\pi_N \colon M \times N \to N$. Mostre que se tem

$$\int_{M\times N} \pi_M^*\alpha \wedge \pi_N^*\beta = \left(\int_M \alpha\right) \cdot \left(\int_N \beta\right).$$

§7.

7.1. Defina uma função $f \colon \mathbb{R}^3 \to \mathbb{R}$ pondo $f(x,y,z) = (\sqrt{x^2+y^2} - 1)^2 + z^2$, mostre que todo número real diferente de zero é valor regular de f e que, se $0 < c < 1$, o conjunto $M = \{(x,y,z) \in \mathbb{R}^3; f(x,y,z) \leq c\}$ é um toro sólido, isto é, uma superfície compacta tri-dimensional, cujo bordo ∂M é um toro de dimensão 2.

7.2. Seja M uma superfície de classe C^k em $\mathbb{R}^n$. Ponha $N = \{(x,v) \in \mathbb{R}^n \times \mathbb{R}^n; x \in M, v \in (T_xM)^\perp, 0 \leq |x| \leq 1\}$. Mostre que N é uma variedade com bordo, de

dimensão n e classe C^{k-1}, cujo bordo é $v_1(M)$, fibrado unitário normal de M. Sejam M compacta e $k \geq 2$. Tome uma vizinhança tubular $V_{2\varepsilon}(M)$ de M em $\mathbb{R}^n$. Mostre que N é difeomorfa a $V_\varepsilon(M)$.

§8.

8.1. Seja ω uma m-forma de classe C^1 num aberto $U \subset \mathbb{R}^{m+1}$. Dado um ponto $x \in U$, indiquemos com $S(r)$ a esfera de centro x e raio $r > 0$, e com $B(r)$ a bola correspondente. Prove que, se $\{e_1, \ldots, e_{m+1}\}$ é a base canônica de $\mathbb{R}^{m+1}$ então vale

$$d\omega(x)(e_1, \ldots, e_{m+1}) = \lim_{r \to 0} \frac{1}{\text{vol. } B(r)} \int_{S(r)} \omega.$$

Conclua que ω é fechada se, e somente se, para todo domínio compacto $D \subset U$ com bordo ∂D de classe C^2 tem-se $\int_{\partial D} \omega = 0$.

8.2. Sejam $F \colon U \to \mathbb{R}^3$ um campo de vetores C^∞ no aberto $U \subset \mathbb{R}^3$ e $M \subset U$ uma superfície C^∞ tal que $F(p) = 0$ para todo $p \in M$. Prove que, para todo $p \in M$, tem-se $\text{rot.}\, F(p) \in T_pM$.

8.3. Seja $D \subset \mathbb{R}^2$ um domínio compacto com fronteira regular. Dada $F \colon D \to \mathbb{R}^2$ de classe C^2, com $F(x) = (f(x), g(x))$ tem-se $\int_D \text{det. } F'(x)dx = \int_{\partial D} f\, dg$.

8.4. Seja M uma superfície orientada. Se um difeomorfismo $f \colon M \to M$ é C^2-homotópico propriamente à identidade então f preserva orientação.

8.5. Seja ω uma forma de classe C^∞ e grau 2 em $\mathbb{R}^3$, tal que $d\omega \neq 0$ em todos os pontos da bola unitária aberta de $\mathbb{R}^3$. Mostre que existe algum ponto x_0 da esfera unitária de $\mathbb{R}^3$ com $\omega(x_0) \neq 0$.

8.6. Seja $f \colon U \to \mathbb{R}^m$ de classe C^2 no aberto $U \subset \mathbb{R}^m$. Indique com $J(x) = \text{det. } f'(x)$ o determinante jacobiano de f no ponto $x \in U$. Suponha que $D \subset U$ é um domínio compacto com fronteira ∂D de classe C^2. Se $f(x) = 0$ para todo $x \in \partial D$, mostre que $\int_D J(x)dx = 0$.

8.7. Se m é par, então toda forma contínua de grau 1 sobre a esfera S^m deve anular-se em algum ponto.

8.8. Se m é par, toda forma contínua de grau $m-1$ sobre a esfera S^m deve anular-se em algum ponto.

8.9. Seja $v \colon S^m \to \mathbb{R}^{m+1}$ um campo contínuo de vetores (não necessariamente tangentes) na esfera S^m. Se m é par, existe pelo menos um ponto $x \in S^m$ tal que $v(x) = \lambda \cdot x$, com $\lambda \in \mathbb{R}$.

8.10. Uma função $f \colon U \to \mathbb{R}$, de classe C^2 num aberto $U \subset \mathbb{R}^m$, chama-se *harmônica* quando $\Delta f = \Sigma \dfrac{\partial^2 f}{\partial x_i^2} = 0$ em todos os pontos de U. Prove que f é harmônica se, e somente se, para todo domínio compacto com fronteira regular $D \subset U$ tem-se $\displaystyle\int_{\partial M} \frac{\partial f}{\partial N} = 0$ onde $\dfrac{\partial f}{\partial N} \colon \partial M \to \mathbb{R}$ é a derivada de f na direção de N, normal unitária exterior a ∂M.

8.11. Sejam $u, v \colon U \to \mathbb{R}$ funções de classe C^2 no aberto $U \subset \mathbb{R}^m$. Seja $D \subset U$ um domínio compacto com fronteira regular. Aplicando o Teorema da Divergência

$\int_D \text{div.}\, F = \int_{\partial D} \langle F, N\rangle$ ao campo de vetores $F = u\cdot\text{grad.}\, v$ obtenha a *primeira fórmula de Green*:

$$\int_D [u\Delta v + \langle\, \text{grad.}\, u, \text{grad.}\, v\rangle]dx = \int_{\partial D} u\, \frac{\partial v}{\partial N}\cdot$$

Usando esta fórmula com $u = v$, conclua que se u é uma função harmônica que se anula no bordo ∂D então u é identicamente nula em D. Noutras palavras: duas funções harmônicas que coincidam no bordo ∂D coincidem em todo o domínio D.

8.12. Obtenha a *segunda fórmula de Green*:

$$\int_D (u\cdot\Delta v - v\cdot\Delta u)dx = \int_{\partial D}\left(u\,\frac{\partial v}{\partial N} - v\,\frac{\partial u}{\partial N}\right)$$

nas mesmas condições da primeira fórmula.

8.13. Seja D um domínio compacto, conexo, com fronteira regular em $\mathbb{R}^m$. Duas funções harmônicas u, v, (definidas num aberto contendo D) que possuam as mesmas derivadas normais $\dfrac{\partial u}{\partial N} = \dfrac{\partial v}{\partial N}$ no bordo ∂D, diferem por uma constante em D. Em particular, se u é harmônica e $\dfrac{\partial u}{\partial N} = 0$ em ∂D então u é constante em D. [Isto significa que se D é um sólido termicamente isolado, a única distribuição estacionária de temperatura em D é a constante.]

8.14. Nas condições do exercício anterior, se $\dfrac{\partial u}{\partial N} \geq 0$ em ∂D então u é constante em D.

8.15. Se u é harmônica num domínio compacto $D \subset \mathbb{R}^3$ com fronteira regular $\partial D = S$ e x_0 é qualquer ponto de D então

$$u(x_0) = \frac{1}{4\pi}\int_S \left[\frac{1}{|x - x_0|}\,\frac{\partial u}{\partial n} - u\cdot\frac{\partial}{\partial n}\,\frac{1}{|x - x_0|}\right] dS.$$

$\Bigg[$*Sugestão*: Use 8.12 com $v(x) = \dfrac{1}{|x - x_0|}$ no domínio $D - B(x_0;\varepsilon)$, e faça $\varepsilon \to 0.\Bigg]$

8.16. Seja $A \subset \mathbb{R}^m$ um domínio não-compacto, cujo bordo é uma hiperfície compacta ∂A. Seja $u\colon A \to \mathbb{R}$ uma função harmônica tal que $|x|\cdot u(x)$ e $|x|^3\cdot\text{grad.}\, u(x)$ são limitadas em A. Prove que se u (ou $\partial u/\partial N$) se anula em ∂A então u é identicamente nula.

8.17. Seja D um domínio compacto, simplesmente conexo, com fronteira regular. Um campo de vetores de classe C^1 em D, com rotacional e divergência nulos, é o gradiente de uma função harmônica em D. Conclua que dois campos de classe C^1 em D, que têm a mesma divergência, o mesmo rotacional e iguais projeções sobre a normal exterior a ∂D, são iguais.

8.18. Sejam $u, v\colon U \to \mathbb{R}$ harmônicas no aberto $U \subset \mathbb{R}^m$, no qual está contido um domínio compacto conexo D, com fronteira regular ∂D. Seja $h\colon \partial D \to \mathbb{R}$ uma função contínua não-negativa. Se $\dfrac{\partial u}{\partial N} + hu = \dfrac{\partial v}{\partial N} + hv$ em ∂D então $u = v$ em D.

8.19. Seja $B \subset \mathbb{R}^m$ uma bola fechada com centro 0. Mostre que a aplicação de inclusão $i\colon \mathbb{R}^m - B \to \mathbb{R}^m - \{0\}$ induz isomorfismos $i^*\colon H^k(\mathbb{R}^m - B) \to H^k(\mathbb{R}^m - \{0\})$ para $k = 0, 1, 2, \ldots$.

8.20. Para todo $m \geq 2$, considere as transformações lineares $A_m\colon H^{m-1}(\mathbb{R}^m - \{0\}) \to \mathbb{R}$, $B_m\colon H_c^m(\mathbb{R}^m) \to \mathbb{R}$, $C_m\colon H^m(S^m) \to \mathbb{R}$, definidas por $A_m \cdot [\omega] = \int_S \omega$, onde S é qualquer esfera de centro 0 em $\mathbb{R}^m$, $B_m \cdot [\omega] = \int_{\mathbb{R}^m} \omega$ e $C_m \cdot [\omega] = \int_{S^m} \omega$. Prove que cada uma das afirmações abaixo implica a seguinte:

1) A_m é um isomorfismo; 3) C_m é um isomorfismo;
2) B_m é um isomorfismo; 4) A_{m+1} é um isomorfismo.

Daí conclua, por indução, que todas as afirmações acima são verdadeiras, qualquer que seja m.
[*Sugestão*: Para provar que (1) $\Rightarrow$ (2), seja ω uma m-forma C^∞ com suporte compacto em $\mathbb{R}^m$, tal que $d\omega = 0$ e $\int_{\mathbb{R}^m} \omega = 0$. Pelo Lema de Poincaré, existe α de grau $m-1$ em $\mathbb{R}^m$, com $d\alpha = \omega$. A forma α pode não ter suporte compacto. Seja, porém, B uma bola fechada de centro 0, cujo interior contém o suporte de ω. Ponha $S = \partial B$. Pelo Teorema de Stokes, $\int_S \alpha = 0$. Segue-se então de (1) que existe uma forma β, de grau $m-2$, em $\mathbb{R}^m - B$, tal que $d\beta = \alpha$ no complementar de B. Sejam D uma bola de centro 0, contendo B em seu interior, e $\xi\colon \mathbb{R}^m \to \mathbb{R}$ uma função C^∞, igual a 1 em $\mathbb{R}^m - D$ e igual a zero em B. Tome $\gamma = \xi \cdot \beta$ em $\mathbb{R}^m - B$ e $\gamma = 0$ em B. Então $\varphi = \alpha - d\gamma$ tem suporte compacto e é tal que $d\varphi = \omega$.]

8.21. Admita que a superfície $N \subset \mathbb{R}^p$, de classe C^k, possui uma vizinhança aberta V, na qual está definida uma retração $\pi\colon V \to N$, de classe C^k. Mostre que toda aplicação contínua $f\colon M \to N$, definida numa superfície M de classe C^k, é homotópica a uma aplicação $g\colon M \to N$, de classe C^k. (*Sugestão*: Para cada $x \in M$, seja $\varepsilon(x) = \text{dist}\,(f(x), \mathbb{R}^p - V)$. Obtenha $g\colon M \to N$ de classe C^k, tal que $|g(x) - f(x)| < \varepsilon(x)$ para todo $x \in M$ e defina $H\colon I \times M \to N$ pondo $H(t,x) = \pi(1-t)f(x) + tg(x)$.) Mostre ainda que, neste caso, se duas aplicações $f, g\colon M \to N$, de classe C^k, são homotópicas então elas são C^k-homotópicas.

8.22. Ainda admitindo (para M e N) a hipótese feita no exercício anterior mostre que se M e N têm o mesmo tipo de homotopia C^0 então elas têm o mesmo tipo de homotopia C^k. Conclua que se duas superfícies C^∞ são homeomorfas então seus grupos de cohomologia (respect. grupos de cohomologia com suportes compactos) são isomorfos.

8.23. Se $m = 2k-1$ é ímpar, mostre que a aplicação $v\colon S^m \to \mathbb{R}^{m+1}$, dada por

$$v(x_1, \ldots, x_{m+1}) = v(x_1, \ldots, x_{2k}) = (-x_2, x_1, \ldots, -x_{2k}, x_{2k-1}),$$

é um campo de vetores tangentes a S^m, de classe C^∞, sem singularidades.

§9.

9.1. Um homeomorfismo local $f\colon M \to N$ é uma aplicação própria se, e somente se, é uma aplicação fechada.

9.2. Todo difeomorfismo de classe C^k de uma superfície conexa é C^k-isotópico a um difeomorfismo que deixa fixo um ponto pré-determinado em M.

9.3. Se m é ímpar e $f\colon S^m \to N$ é uma aplicação de classe C^1 (tomando valores numa superfície conexa orientada de dimensão m) tal que $f(-x) = f(x)$ então o grau de f é um número par.

9.4. Todo difeomorfismo C^∞ que preserva a orientação de $\mathbb{R}^m$ é C^∞-isotópico à identidade. [Reduza ao caso de um difeomorfismo $f\colon \mathbb{R}^m \to \mathbb{R}^m$ tal que $f(0) = 0$. Defina uma isotopia $H\colon I \times \mathbb{R}^m \to \mathbb{R}^m$ entre f e sua derivada $f'(0)$ pondo $H(t,x) = f(tx)/t$ se $t \neq 0$ e $H(0,x) = f'(0)\cdot x$. Em seguida, estabeleça uma isotopia entre $f'(0)$ e a identidade, usando o Apêndice à seção 10.]

§10.

10.1. Seja ω uma forma contínua de grau 1 em $\mathbb{R}^3$ tal que $\omega(x)\cdot v = 0$ sempre que $|x| = 1$ e $\langle v, x\rangle = 0$. Prove que existe $x \in \mathbb{R}^3$ tal que $|x| \leq 1$ e $\omega(x) = 0$.

10.2. Seja $D \subset \mathbb{R}^{m+1}$ um domínio compacto com fronteira regular. Suponha dada uma aplicação $F\colon D \to \mathbb{R}^{m+1}$ de classe C^1, com $F(\partial D) \subset \mathbb{R}^{m+1} - \{0\}$. Dado arbitrariamente $\varepsilon > 0$, mostre que existe $G\colon D \to \mathbb{R}^{m+1}$ de classe C^1 tal que $0 \notin G(\partial D)$, $|G(x) - F(x)| < \varepsilon$, $|G'(x) - F'(x)| < \varepsilon$ para todo $x \in D$ e todos os zeros de G são simples.

10.3. Sejam $F\colon U \to \mathbb{R}^m$ contínua, de classe C^2 no aberto conexo limitado U, $\xi\colon V \to U$ um homeomorfismo tal que $\xi|V$ é um difeomorfismo de classe C^2 do aberto V sobre U e $\lambda\colon \mathbb{R}^m \to \mathbb{R}^m$ um difeomorfismo, ainda de classe C^2. Prove que $gr(\lambda\circ F\circ\xi, V, \lambda(y)) = \pm gr(F, U, y)$ para todo $y \notin F(\partial U)$, onde o sinal é $+$ se, e somente se os determinantes jacobianos de ξ e λ têm o mesmo sinal.

10.4. Seja v um campo de vetores tangentes a uma superfície $M \subset \mathbb{R}^n$. Uma *singularidade* de v é um ponto $x \in M$ tal que $v(x) = 0$. Seja $a \in M$ uma singularidade *isolada* de um campo v, de classe C^2 na superfície M, de classe C^3. (Isto é, $v(x) \neq 0$ para todo $x \neq a$ suficientemente próximo de a.) Tome uma parametrização $\varphi\colon V_0 \to V$ de classe C^3, que se estenda a um homeomorfismo de V_0 sobre V, onde $V \subset \mathbb{R}^m$ é uma bola de centro na origem, e tal que $a \in V$ seja a única singularidade de v em $\overline{V}$. Defina o *índice* da singularidade a pondo $\operatorname{ind}(v;a) = gr(F, V_0, 0)$, onde $F\colon \overline{V}_0 \to \mathbb{R}^m$ é definida por $F(x) = \varphi'(x)^{-1}\cdot v(\varphi(x))$. [Note que, para todo $x \in V_0$, $\varphi'(x)$ é um isomorfismo de $\mathbb{R}^m$ sobre $T_{\varphi(x)}M$.] Se $\psi\colon W_0 \to W$ for outra parametrização do tipo acima, o índice terá uma nova expressão $\operatorname{ind}(v;a) = gr(G, W_0, 0)$, onde, para todo $y \in W_0$, tem-se $G(y) = \xi'(y)^{-1}\cdot F(\xi(y))$, sendo $\xi = \psi^{-1}\circ\varphi$ a mudança de parametrização. Obtenha uma homotopia entre G e a aplicação G_0, dada por $G_0(y) = \xi'(0)^{-1}\cdot F(\xi(y))$, e conclua que o índice da singularidade a não depende da parametrização usada para definí-lo.

10.5. Dada uma superfície compacta $M \subset \mathbb{R}^n$, de classe C^∞, seja $\varepsilon > 0$ tal que dois segmentos de reta fechados, de comprimentos $\leq \varepsilon$, normais a M em dois pontos distintos quaisquer de M sejam disjuntos. O conjunto N, reunião de todos esses segmentos de reta fechados, normais a M em todos os seus pontos é uma superfície compacta n-dimensional com bordo. Seja $\pi\colon N \to M$ a retração C^∞ que associa a cada ponto $y \in N$ o pé do único segmento normal a M que o contém. Dado um campo de vetores tangentes a M, de classe C^∞, com

um número finito de singularidades, estenda-o a uma aplicação $F\colon N \to \mathbb{R}^n$, de classe C^∞, cujos zeros são precisamente as singularidades de v e tal que, para todo $y \in \partial N$, o vetor $F(y)$ aponta para fora de N. [Ponha $F(y) = v(\pi(y)) + (y - \pi(y))$.] Mostre que $gr(F, N, 0)$ é igual ao grau da aplicação normal de Gauss $\gamma\colon \partial N \to S^{n-1}$. Por outro lado, mostre que $gr(F, N, 0)$ é igual à soma dos índices das singularidades de v em M. [Para ver isto, considere, em torno de cada singularidade, uma pequena bola B_i que não contenha outros zeros de F. Então $N - \cup B_i$ é uma superfície com bordo e $gr(F, N - \cup B_i, 0) = 0$.] Conclua que a soma dos índices das singularidades de v não depende do campo v. Tal soma chama-se a *característica de Euler* da superfície M.

10.6. Mostre que se $M \subset \mathbb{R}^{m+1}$ é uma hiperfície (orientada) compacta com m par então o grau da aplicação normal de Gauss $\gamma\colon M \to S^m$ é igual à metade da característica de Euler de M. [Considere N como no exercício anterior e observe que ∂N consiste em duas cópias de N tais que, em cada uma delas a retração $\pi\colon \partial N \to M$ é um difeomorfismo que preserva orientação.]

10.7. Defina, na esfera, S^m, o campo de vetores "norte-sul" pondo, para cada $x \in S^m$, $v(x)$ = projeção ortogonal do vetor $-e_{m+1} = (0, \ldots, 0, -1)$ sobre o plano tangente a S^m no ponto x. Verifique que v é de classe C^∞, possui duas singularidades uma de índice 1, outra de índice $(-1)^m$, e conclua que a característica de Euler de S^m é zero quando m é ímpar e igual a 2 quando m é par.

Bibliografia

[1] Do Carmo, Manfredo P. - *Differential Geometry of Curves and Surfaces*, Prentice Hall, Englewood, Cliffs, N. Jersey, 1976.

[2] Courant, Richard - *Differential and Integral Calculus*, vol. II (Translated by E. J. McShane), Interscience, New York, 1937.

[3] Dieudonné, Jean - *Foundations of Modern Analysis*, Academic Press, New York, 1960.

[4] Dugundji, James - *Topology*, Allyn and Bacon, Boston, 1965.

[5] Graves, Lawrence M. - *The Theory of Functions of Real Variables*, McGraw-Hill Book Company, New York, 1946.

[6] Jordan, Camile - *Cours d'Analyse*, Tome Prémier (Nouvelle Tirage) Gauthiers Villars, Paris, 1959.

[7] Lang, Serge - *Analysis I*, Addison-Wesley, Reading, Mass. 1968.

[8] Lima, Elon Lages - *Espaços Métricos*, Projeto Euclides, IMPA, Rio de Janeiro, 1977.

[9] Lima, Elon Lages - *Análise no Espaço* $\mathbb{R}^n$, Coleção Matemática Universitária, IMPA, 2002.

[10] Lima, Elon Lages - *Grupo Fundamental e Espaços de Recobrimento*, Coleção Matemática Universitária, IMPA, 1998.

[11] Lima, Elon Lages - *Álgebra Exterior*, Coleção Matemática Universitária, IMPA, 2005.

[12] Lima, Elon Lages - *Variedades Diferenciáveis*, Monografias de Matemática, nº 15, IMPA, Rio de Janeiro, 1973.

[13] Lima, Elon Lages - *Introdução à Topologia Diferencial*, Notas de Matemática, nº 23, IMPA, Rio de Janeiro, 1961.

[14] Lima, Elon Lages - *Duas novas demonstrações do Teorema de Jordan-Brouwer no caso diferenciável*, Matemática Universitária, Número 4 (1986), pags. 89-105.

[15] Lloyd, Noel G. - *Degree Theory*, Cambridge Tracts in Mathematics, 73, Cambridge University Press (1978).

[16] Milnor, John - *Topology from the Differentiable Viewpoint*, Charlottesville, The University Press of Virginia, 1965.

[17] Milnor, John - *Analytic proofs of the "hairy ball theorem" and the Brouwer fixed point theorem*, American Mathematical Monthly, vol. 85 (1978), pags. 521-524.

[18] Munkres, James - *Topology, a First Course*, Prentice Hall, Englewood Cliffs, N. Jersey, 1975.

[19] Nirenberg, Louis - *Topics in Non Linear Functional Analysis*, New York University, 1974.

[20] Phillips, E. G. - *A Course of Analysis*, Cambridge University Press, Cambridge, 1956.

[21] Picard, Émile - *Traité d'Analyse*, Tome 1, Gauthier Villars, Paris, 1891.

[22] Radó, Tibor - *Lectures on the Plateau Problem*, Chelsea Publishing Company, New York, 1951.

[23] De Rham, Georges - *La théorie des formes différentielles extérieures et l'homologie des variétés différentiables*, Rendiconti di Matematica e delle sue Applicazioni, vol. 20 (1961), pags. 105-146.

[24] Rogers, C. A. - *A less strange version of Milnor's proof of Brouwer's fixed point theorem*, American Mathematical Monthly, vol. 87 (1980), pags. 525-527.

[25] Rudin, Walter - *Princípios de Análise Matemática* (Tradução de Eliana Rocha Henriques de Brito), Ao Livro Técnico, 1971.

[26] Scheeffer, Ludwig - *Allgemeine untersuchungen über rectification der Curven*, Acta Mathematica, vol. 5 (1884), pags. 49-82.

[27] Schwartz, Jacob T. - *The formula for change in variables in a multiple integral*, American Mathematical Monthly, vol. 61 (1954), pags. 81-85.

[28] Schwarz, Hermann A. - *Sur une définition erronée de l'aire d'une surface courbe*, Gesammelte Mathematische Abhandlungen (Zweiter Band), pags. 309-311, Berlin, 1890.

[29] Sieberg, Hans W. - *Some historical remarks concerning degree theory*, American Mathematical Monthly, vol. 88 (1981), pags. 125-139.

[30] Spivak, Michael - *A Comprehensive Introduction to Differential Geometry*, volume one (second edition), Publish or Perish, Berkeley, 1979.

[31] Spivak, Michael - *Calculus on Manifolds*, Benjamim, New York, 1965.

[32] Valiron, Georges - *Théorie des Fonctions* (2.me édition), Masson, Paris, 1947.

Índice Remissivo

Índice Remissivo

elon lages lima

Elon Lages Lima é Pesquisador Emérito do Instituto Nacional de Matemática Pura e Aplicada (IMPA), Doutor Honoris Causa pela Universidade Federal de Alagoas, pela Universidade Federal do Amazonas e pela Universidad Nacional de Ingenieria del Perú, professor Honoris Causa da Universidade Federal do Ceará, da Universidade Federal da Bahia, da Universidade Estadual de Campinas e da Pontifícia Universidade Católica do Perú, além de membro titular da Academia Brasileira de Ciências e da Academy of Sciences for the Developing World. É autor de vários livros de Topologia, Análise, Álgebra e Matemática Elementar, dois dos quais são ganhadores do Prêmio Jabuti.

curso de análise vol. 2

Instituto Nacional de Matemática Pura e Aplicada